КИЇВСЬКИЙ ІНСТИТУТ БАНКІВСЬКОЇ СПРАВИ

ДАНИЛОВ В. Я.
ДАНІЛОВ В. Я.
ПАСЕНЧЕНКО Ю. А.

ТЕОРІЯ ЙМОВІРНОСТЕЙ

навчальний посібник
для студентів вищих закладів

Київ – 2014

Укладачі: ДАНИЛОВ Валерій Якович, доктор техн. наук, професор
ДАНІЛОВ Володимир Якович, канд. фіз.-мат. наук, доцент
ПАСЕНЧЕНКО Юрій Антонович, канд. фіз.-мат. наук, доцент

Розглянуто та схвалено на засіданні кафедри економічної теорії Київського інституту банківської справи "16" квітня 2014, протокол № 8

Затверджено методичною радою Київського інституту банківської справи " 15 " травня 2014, протокол № 9

Рецензенти: Новіцький В. В., доктор фізико-математичних наук, професор, інститут математики НАН України;
Кушніренко С. В., кандидат фізико-математичних наук, доцент, Київський національний університет ім. Тараса Шевченка.

Навчальний посібник призначений для вивчення дисципліни «Теорія ймовірностей та математична статистика» студентами вузів, що готують банківських спеціалістів. Він містить методичні рекомендації та завдання до розділів: випадкові події, випадкові величини, елементи математичної статистики. Перші два розділи охоплюють основні поняття теорії ймовірностей. В розділі елементи математичної статистики вивчається, як методи теорії ймовірностей застосовуються до відбору та обробки спостережень при вивченні ймовірнісних характеристик випадкових явищ.Наведені у посібнику приклади ілюструють застосування методів теорії ймовірностей та математичної статистики у банківській діяльності.

Посібник може використовуватись студентами економічних спеціальностей вищих навчальних закладів, а також буде корисним усім, хто застосовує ймовірнісні методи на практиці.

Передмова

Процеси, що відбуваються в природі та суспільстві, є наслідком взаємодії багатьох факторів. Для того, щоб вивчити ці процеси і в подальшому керувати ними, слід з'ясувати, яку роль у досліджуваному процесі відіграє кожний фактор окремо. Так, аналізуючи процес зміни курсу національної валюти потрібно з'ясувати вплив ряду економічних та соціальних факторів, які можуть істотно змінювати курс валюти відносно інших валют.

Різноманітні питання математики, фізики, хімії, економіки, соціології та ряду інших галузей знань приводять до необхідності встановлення залежності між величинами, що описують той чи інший еволюційний процес. За формою прояву причинних зв'язків закони природи та суспільства поділяються на два класи – детерміновані та стохастичні.

Теорія ймовірностей вивчає ймовірнісні характеристики випадкових явищ зокрема, ймовірності випадкових подій, функції розподілу випадкових величин та їх числові характеристики (математичне сподівання, дисперсія, середньоквадратичне відхилення, та ін.). Основна властивість будь-якої випадкової події незалежно від її природи – ймовірність її здійснення.

Теорія ймовірностей є основою для математичної статистики, основним змістом якої є систематизація, обробка і подальше використання статистичної інформації для встановлення статистичних закономірностей певної ознаки або ряду ознак сукупності елементів.

Використовуючи вибірковий метод, математична статистика по обмеженим даним (вибірці) відновлює з певним степенем достовірності характеристики, які притаманні генеральній сукупності, тобто всьому можливому набору даних, що описують досліджуване явище. Математична статистика дозволяє здійснювати обчислення наближених значень ймовірносних характеристик. Найтиповішими задачами математичної статистики є оцінка невідомої функції розподілу, оцінка невідомих параметрів розподілу, перевірка статистичних гіпотез.

Метою даного посібника є вивчення основ теорії ймовірностей та математичної статистики, набуття навиків розв'язування типових ймовірносних задач та застосування їх для здійснення математичного аналізу стохастичного експерименту і побудови його математичної моделі. Посібник складається з трьох розділів: випадкові події, випадкові величини та елементи математичної статистики. Розділи випадкові події та випадкові величини охоплюють основні поняття теорії ймовірностей. В розділі елементи математичної статистики вивчаються методи відбору та обробки спостережень випадкових явищ для виявлення їх ймовірносних характеристик.

Даний посібник розроблено у відповідності до типової програми навчальної дисципліни "Теорія ймовірностей і математична статистика" для студентів економічних спеціальностей вищих навчальних закладів.

Розділ 1. Випадкові події

Глава 1. Основні поняття теорії ймовірностей

§1. Предмет теорії ймовірностей

Для широкого кола явищ спостерігається неоднозначність наслідків (результатів) досліду при його повторенні в одних і тих же умовах. До них, наприклад, відносяться різноманітні ігри, вибірковий контроль якості (стандартності) товарів, прибуток підприємства, врожайність зернових, щоденний видобуток вугілля і т.п. Абстрагуючись від конкретного змісту цих явищ, можна говорити про деякий експеримент, результат якого не можна заздалегідь передбачити і його можна повторювати будь-яку кількість разів. Відмітимо, що при повторенні експерименту за одних і тих самих умов результати можуть бути різними. Такий експеримент називають ***випадковим експериментом (випадковим випробуванням).***

Означення. Подія, яка при виконанні певного комплексу умов у данному експерименті може відбутися або не відбутися, називається ***випадковою.*** Події „ при підкиданні монети з'явиться цифра"; „ при підкиданні грального кубика (це кубик з однорідного матеріалу, на гранях якого вибито цифри від одного до шести) випадає грань з числом п'ять"; „вибране навмання підприємство має 40% прибутку"; „ вибраний навмання працівник банку має вищу банківську освіту" є випадковими.

Означення. Подія, яка напевно відбудеться у даному випробуванні, здійснюваного з додержанням певного комплексу умов, називається **вірогідною** подією. Подія "при одному підкиданні грального кубика на верхній грані буде довільне ціле число від одного до шести" є вірогідною.

Означення. Подія, яка напевно не відбудеться ніколи при виконанні певного комплексу умов, називається ***неможливою.*** Подія "при одному підкиданні грального кубика на верхній грані буде число очок, що дорівнює сім" є неможливою.

Теорія ймовірностей не може передбачати, чи відбудеться випадкова подія чи ні. Але, коли розглядаються випадкові події, які можемо спостерігати багатократно при виконанні одних і тих же умов, то вони підпорядковані певним закономірностям, і вивчення цих закономірностей при допомозі математичних обчислень є предметом теорії ймовірностей.

§2. Простори елементарних подій. Операції над подіями

Основним методом математичного опису випадкового явища є побудова для нього математичної моделі, яка б правильно відображала всі суттєві його сторони і одночасно була б простою з математичної точки зору. Відправною точкою при побудові математичної моделі є поняття елементарних подій, що пов'язані явищами, що вивчаються. Під ними розуміють множину взаємовиключних наслідків (результатів) експерименту, тобто таких, що в результаті проведення експерименту (випробування) настає один і тільки один з них. Цю множину наслідків (результатів) експерименту назвемо ***простором елементарних подій*** і позначимо через Ω.

Зауважимо, що побудова простору елементарних подій є процедурою неоднозначною, і тому Ω вибирається найбільш вигідним чином. Простір елементарних подій може мати різну природу – бути скінченним, зліченним або неперервним.

Наприклад, якщо підкинути гральний кубик один раз, то наслідками експерименту (випробування) можуть бути події w_i $\left(i=1,2,3,4,5,6\right)$, які полягають в тому, що на верхній грані кубика випало i очок. Простір елементарних подій $\Omega=\left\{w_1,w_2,w_3,w_4,w_5,w_6\right\}$ є скінченний. Надалі, через $|A|$ позначатимемо кількість елементів множини A.

Наведемо приклад зчисленного простору подій. Монету підкидають до першого випадання герба. Позначимо через $w_i=\left\{Ц,Ц,\ldots,Ц,Г\right\}$ послідовність підкидань монети до першого випадання герба (перші $i-1$ раз випала цифра, а потім випав герб). Тоді простір елементарних подій $\Omega=\left\{w_1,w_2,\ldots,w_i,\ldots\right\}$ може бути множина нескінченна, але зліченна (елементи її можна занумерувати).

Прикладом нескінченного і незліченного простору елементарних подій може служити множина наслідків при "киданні" точки на відрізок або на круг. В результаті проведенного випробування випадковим наслідком є попадання в будь – яку точку відрізка або круга і множина цих наслідків є нескінченна, при цьому перенумерувати їх неможливо.

Виходячи з поняття простору Ω, довільна подія математично визначається наступним чином.

Означення. Якщо Ω - дискретний простір (скінченний або зчисленний), то ***випадковою подією*** або просто ***подією*** називається будь-яка підмножина множини Ω.

Випадкові події позначатимемо літерами $A,B,C,H,\ldots$. Будь – яка випадкова подія $A\subset\Omega$ (A - підмножина Ω) складається з елементарних подій.

Неможлива подія, тобто подія, яка в результаті експерименту не може відбутися, ототожнюється з порожньою множиною $\emptyset$.

Приклад 1. При підкиданні грального кубика один раз вказати підмножину елементарних подій, яка утворює події: $A=${на верхній грані куба випадає парне число очок}; $B=${ на верхній грані куба випадає число очок, яке менше трьох}; $C=${ на верхній грані куба випадає дробове число очок}.

Розв'язування. Простір елементарних подій $\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6\}$, де елементарна подія w_i, $(i = 1,2,3,4,5,6)$ полягає в тому, чи при підкиданні грального кубика випадає число очок, рівне i. Тепер маємо, що $A = \{w_2, w_4, w_6\}$, $B = \{w_1, w_2\}$, $C = \emptyset$ - порожня множина. Очевидно, множини A, B і C є підмножинами множини Ω.

Оскільки останнє визначення події ототожнює її з множиною, то над подіями можна виконувати всі операції, які виконуються над множинами.

Означення. ***Сумою двох подій*** A ***і*** B ***називається подія*** $A \cup B$ ***(або*** $A + B$***), яка складається з усіх елементарних подій, що належать*** A ***або*** B.

В теорії ймовірностей сума подій $A \cup B$ означає, що відбувається принаймні одна з подій A або B в результаті експерименту.

В загальному випадку, під сумою декількох подій розуміють подію, яка полягає в тому, що хоча б одна з цих подій відбудеться.

Означення . ***Добутком двох подій*** A ***і*** B ***називається подія*** $A \cap B$ ***(або*** AB***), яка складається з усіх елементарних подій , які належать*** A ***і*** B ***одночасно .***

В теорії ймовірностей подія $A \cap B$відбувається тоді і лише тоді, коли відбуваються події A і B.

В загальному випадку, добутком декількох подій називають подію, яка полягає в тому, що всі події одночасно відбудуться.

Означення. ***Різницею двох подій*** A ***і*** B ***називається подія*** $A \setminus B$***, яка складається з елементарних подій, які належать*** A ***і не належать*** B.

З точки зору теорії ймовірностей подія $A \setminus B$, полягає в тому, що подія A відбувається, а подія B не відбувається.

Означення. ***Подія*** $\overline{A} = \Omega \setminus A$ ***називається протилежною до події*** A. Подія $\overline{A}$ означає, що подія A не відбувається.

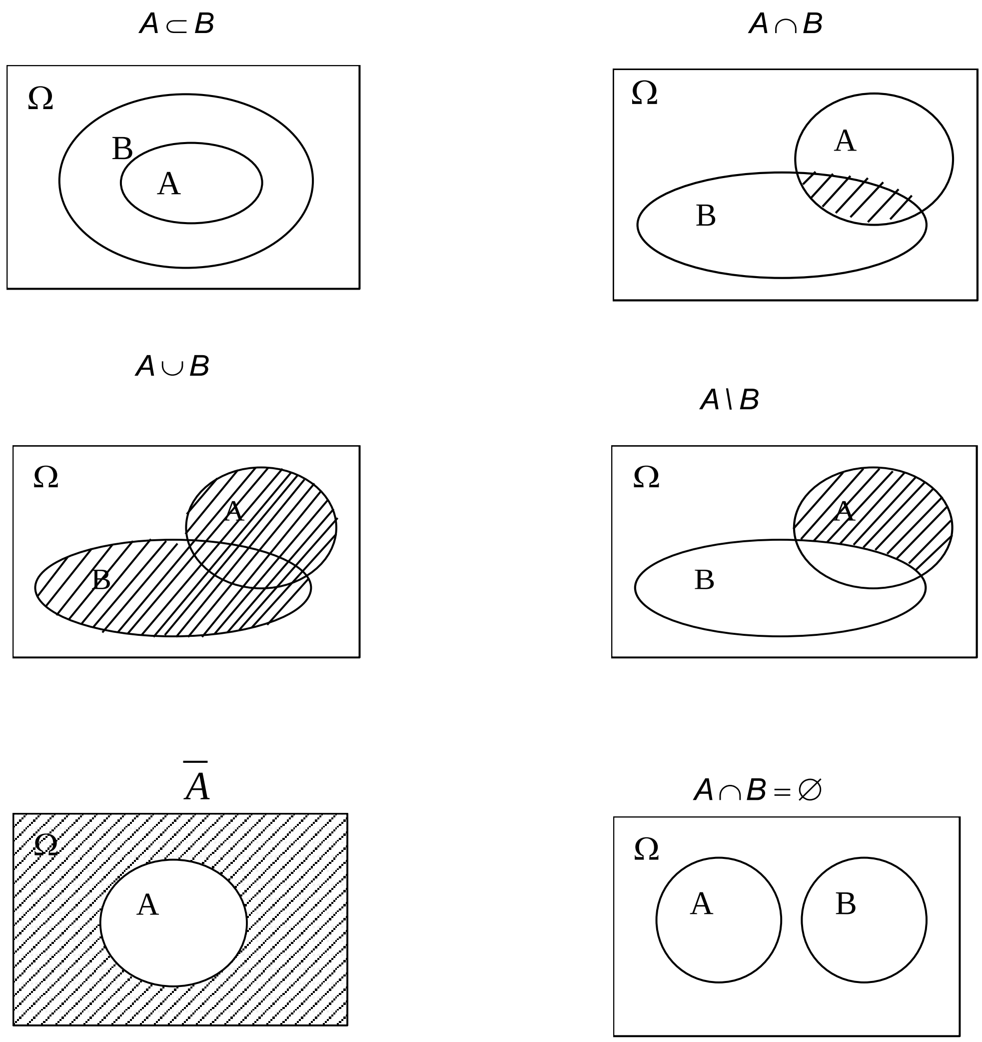
Означення. ***Події*** A ***і*** B ***називаються несумісними, якщо*** $AB = \emptyset$ ***(порожня множина).***

З позиції теорії ймовірностей це означає, що поява однієї події виключає появу іншої в одному і тому ж випробуванні.

Означення. ***Якщо*** $A \subset B$ ***і*** $B \subset A$***, то події називаються рівносильними і записують*** $A = B$.

Зауваження. Поняття суми і добутку подій можна узагальнити на нескінченне число подій. При цьому використовують позначення: якщо $A_n \subset \Omega$, $n = 1,2,3,\ldots$, то $\bigcup_{n=1}^{\infty} A_n$ і $\bigcap_{n=1}^{\infty} A_n$ позначають відповідно суму і добуток нескінченного числа подій.

Дії над подіями можна геометрично інтерпретувати діаграмами Ейлера-Венна мал.1. (подіям $A \cup B$, $A \cap B$, $A \subset B$, $A \setminus B$, $\overline{A}$ відповідають зафарбовані області).



Мал. 1. Діаграми Ейлера-Венна

В силу теоретико-множинної інтерпретації подій, операції над подіями мають тіж властивості, що й операціїї над множинами:

1. $A \cap \Omega = A;$ $A \cap \emptyset = \emptyset;$ $A \cap \Omega = A;$ $A \cup \Omega = \Omega;$ $A \cap \overline{A} = \emptyset;$ $A \cup \overline{A} = \Omega;$ $A \setminus B = A \cap \overline{B};$ $\emptyset \subseteq \mathrm{A} \subseteq \Omega$ для будь-якої множини A.

2. $A \cap B = B \cap A;$ $A \cup B = B \cup A$ (*переставна властивість перетину та об'єднання*).

3. $A \cap (B \cap C) = (A \cap B) \cap C;$ $A \cup (B \cup C) = (A \cup B) \cup C$ (*сполучна властивість перетину та об'єднання*).

4. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (*розподільна властивість перетину відносно об'єднання*).

5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$(*розподільна властивість об'єднання відносно перетину*)

6. $\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{A \cap B} = \overline{A} \cup \overline{B}$ (*формули двоїстості*).

Приклад 2. Нехай $A,\ B$ і C три довільних події. Описати події, які полягають в тім, що:

а) відбулися всі три події;
б) відбулася рівно одна подія;
в) відбулося рівно дві події;
г) відбулася хоча б одна подія;
д) відбулося не більше двох подій;
е) не відбулося жодної події.

Розв'язування. а) подія $A_1 = \{$ відбулися всі три події$\}$ - $A_1 = A\cap B\cap C$; б) подія $A_2 = \{$відбулася рівно одна подія$\}$ - $A_2 = (A\cap \overline{B}\cap \overline{C})\cup(\overline{A}\cap B\cap \overline{C})\cup$

$\cup(\overline{A}\cap \overline{B}\cap C)$; в) $A_3 = \{$відбулося рівно дві події$\}$ - $A_3 = (A\cap B\cap \overline{C})$ $)\cup(A\cap \overline{B}\cap C)\cup(\overline{A}\cap B\cap C$; г) $A_4 = \{$відбулася хоча б одна подія$\}$ - $A_4 =$ $(A\cap \overline{B}\cap \overline{C})\cup(\overline{A}\cap B\cap \overline{C})\cup(\overline{A}\cap \overline{B}\cap C)\cup(A\cap \overline{B}\cap C)\cup$

$\cup(\overline{A}\cap B\cap C)\cup(A\cap B\cap \overline{C})\cup(A\cap B\cap C)$; д) $A_5 = \{$відбулося не більше двох подій$\}$ - $A_5 = (A\cap \overline{B}\cap \overline{C})\cup(\overline{A}\cap B\cap \overline{C})\cup(\overline{A}\cap \overline{B}\cap C)\cup$ $(A\cap \overline{B}\cap C)\cup\cup(\overline{A}\cap B\cap C)\cup(A\cap B\cap \overline{C})$;

е) $A_6 = \{$не відбулося жодної події$\}$ - $A_6 = \overline{A}\cap \overline{B}\cap \overline{C}$.

§3. *Класичне визначення ймовірності події*

Ймовірність події – одне з основних понять теорії ймовірностей. Існує декілька визначень ймовірності події і тут розглянемо одне з них, яке називають ***класичним.***

Спочатку пояснимо це поняття на прикладі. Нехай в урні міститься вісім однакових куль, серед яких чотири сині, три жовті і одна біла (непофарбована). З урни навмання виймаємо одну кулю – проводимо випробування. Очевидно можливість вийняти з урни пофарбовану кулю (синю або жовту) є більшою ніж можливість вийняти білу (непофарбовану) кулю. Поставимо перед собою задачу: дати кількісну оцінку можливості вийняти з урни пофарбовану кулю.

Нехай подія $A = \{$вийнята навмання з урни куля є пофарбована$\}$. В нашому прикладі можливі такі елементарні наслідки: w_1, w_2, w_3, w_4 - вийнята синя куля; w_5, w_6, w_7 - вийнята жовта куля; w_8 - вийнята біла куля. Зауважимо, що можливості для кожної з елементарних подій w_i $(i = \overline{1,8})$ є однакові, принаймні одна з них в результаті випробування обов'язково відбудеться, кожна з них єдиноможлива і вони несумісні. При цьому, елементарні наслідки $w_1, w_2, w_3, w_4, w_5, w_6, w_7$ є сприятливі для події A, а елементарний наслідок w_8 - несприятливий для події A.

Звідси випливає, що за число, яке характеризує можливість появи події A, можна прийняти відношення числа сприятливих для цієї події наслідків до числа всіх можливих наслідків випробування. Описане число і називають ймовірністю події A.

Перейдемо до загального визначення ймовірності події. Нехай подія A- деякий наслідок випробування і $w_1, w_2, ..., w_n$ - скінченна множина всіх можливих і єдиноможливих попарно несумісних наслідків цього випробування (повна група елементарних наслідків). Отже, подія A відбувається тоді і тільки тоді, коли відбудуться деякі події із сукупності w_i, $i = \overline{1,n}$, що є сприятливими для події A.

Припустимо, що елементарні події $w_1, w_2, ..., w_n$ ***рівноможливі***, тобто немає підстав вважати, що одна з цих подій є більш можливою, ніж інша.

Означення. ***Ймовірністю $P(A)$ події A називається відношення числа рівноможливих елементарних наслідків, які сприятливі для події A, до числа всіх рівноможливих і єдиноможливих елементарних наслідків випробування.***

Отже, якщо m- число сприятливих для події A наслідків, а n- загальне число всіх елементарних наслідків в даному випробуванні, то

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

З наведеного визначення ймовірності випливають такі її основні властивості:

1. Оскільки $0 \le m \le n$, то $0 \le P(A) \le 1$, тобто ймовірність будь – якої події є невід'ємне число, що не перевищує одиниці.

2. ***Ймовірність неможливої події дорівнює нулю.*** Дійсно, якщо подія A неможлива, то число сприятливих для неї наслідків $m = 0$ і маємо, що

$$P(A) = \frac{0}{n} = 0 .$$

3. ***Ймовірність достовірної події дорівнює одиниці.*** Якщо подія A достовірна, то $m = n$ і, значить,

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

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

Розглянемо приклади обчислення ймовірності за формулою (1).

Приклад 3. Гральний кубик кидають один раз. Знайти ймовірність того, що в результаті випробування випаде грань з непарним числом очок.

Розв'язування. Позначимо подію $A = \{$число очок, що випадуть при киданні гральний кубик, є непарне$\}$. Простір елементарних подій (множина всіх можливих наслідків випробування) $\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6\}$, тобто $n = 6$. Сприятливими для події A наслідками є w_1, w_3, w_5, тобто $m = 3$. Всі елементарні наслідки є рівноможливі, несумісні і утворюють повну групу. Тому шукана ймовірність

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

Приклад 4. По цілі провели 90 пострілів, з яких зареєстровано 60 влучень. Обчислити ймовірність влучення в ціль при одному пострілі.

Розв'язування. Позначимо подію $A=\{$влучення в ціль при одному пострілі$\}$. Тоді множина всіх можливих наслідків $n=90$, а множина сприятливих для події A наслідків - $m=60$. Отже, згідно (1):

$$P(A)=\frac{60}{90}=\frac{2}{3}.$$

Образно висловлюючись, ймовірністю події є міра об'єктивної впевненості в тому, що дана подія відбудеться.

§4. *Елементи комбінаторного аналізу*

З комбінаторними обчисленнями доводиться мати справу представникам багатьох спеціальностей: диспетчеру при складанні графіка руху поїздів, хіміку при розгляді різних можливих типів зв'язків атомів у молекулах, біологу при вивченні різних послідовностей чергування амінокислот в білкових сполуках.

Комбінаторні міркування лежать в основі розв'язання багатьох задач теорії ймовірностей. В будь-якій комбінаторній задачі в тій чи іншій мірі фігурує так зване *правило множення*.

Як приклад розглянемо наступну задачу. З міста A в місто B веде n різних шляхів, а з міста B в місто C веде m різних доріг. Скількома способами можна здійснити подорож з A міста в місто C? Обираючи один з n можливих шляхів з міста A в місто B , ми маємо m різних продовжень подорожі з міста B в місто C. Тому маємо відповідь - $n\cdot m$.

Основне правило комбінаторики. Нехай треба виконати одну за одною k дій. Якщо одну дію можна здійснити n_1 способом, другу дію (після того, як здійснена перша) можна виконати n_2 способами, третю дію (після того, як здійснені попередні дві) -n_3 способами і т.д., k-ту - n_k способами, то послідовність усіх k дій може бути виконана $n_1\cdot n_2\cdot n_3\cdot\ldots\cdot n_k$ способами.

Приклад 5. Скількома способами можна скласти групу делегатів на банківський тренінг, яка б складалась з одного бухгалтера і одного економіста, якщо в групі навчається 15 економістів і 9 бухгалтерів?

Яка імовірність того, що ця делегація буде складатись з відмінників навчання, якщо 6 бухгалтерів і 4 економісти навчаються на "відмінно"?

Розв'язування. Згідно основного правила комбінаторики, делегацію можна вибрати $|\Omega|=n=15\cdot 9=135$ способами.

Нехай подія $A=\{$ обидва делегати – відмінники$\}$. Тоді

$$|A|=m=6\cdot 4=24,\ \ P(A)=\frac{|A|}{|\Omega|}=\frac{m}{n}=\frac{24}{135}=\frac{8}{45}.$$

Число k-елементних підмножин множини X. Виясними питання, скільки серед всіх підмножин множини X є k -елементних підмножин, тобто підмножин, які містять k елементів? Позначимо через C_n^k число k-елементних підмножин множини, яка містить n елементів. Наведемо без доведення наступну теорему (див.наприклад [1]).

Теорема 1. ***Число всіх k - елементних підмножин множини, що складається з n елементів, дорівнює***

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

Символ $n!$ (читається n-факторіал) означає добуток $1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ всіх натуральних чисел від 1 до n включно (прийнято, що $0! =1, 1!=1$).

Приклад 6. Скількома способами з 20 студентів можна обрати делегацію на конференцію Ліги студентів, яка складається з чотирьох студентів?

Розв'язування. Шукане число способів дорівнює числу чотирьохелементних підмножин в множині з 20 елементів:

$$C_{20}^4 = \frac{20!}{4!\cdot(20-4)!} = \frac{20!}{4!\cdot 16!} = \frac{17 \cdot 18 \cdot 19 \cdot 20}{1 \cdot 2 \cdot 3 \cdot 4} = 4845.$$

Приклад 7. Скількома способами з 10 позик банку можна обрати 3 довгострокові позики?

Розв'язання. Шукане число способів дорівнює числу трьохелементних підмножин в множині з семи елементів:

$$C_{10}^3 = \frac{10!}{3!\cdot 7!} = \frac{8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3} = 120.$$

Приклад 8. ***(гіпергеометричний розподіл).*** Портфель позик комерційного банку виробів складається з M позик, з них m довгострокових (тобто $M-m$ короткотермінових). Яка імовірність того, що, вибираючи навмання для аналізу L позик $(L < M)$, можна вибрати k довготермінових позик $(k < m)$?

Розв'язування. Оскільки порядок вибраних позик не має значення, то загальне число способів вибрати L позик з M дорівнює $|\Omega| = n = C_M^L$. Нехай подія B={з L обраних позик k позик є довготермінових} і, отже, $(L-k)$ - короткотермінові. Необхідне число довготермінових позик можна вибрати C_m^k способами, а короткотермінових можна вибрати C_{M-m}^{L-k} способами. Тоді, згідно основного правила комбінаторики, пару позик можна вибрати $|B| = K = C_m^\ell \cdot C_{M-m}^{L-\ell}$ способами.

Отже, шукана імовірність події B дорівнює:

$$P(B) = C_m^\ell \cdot C_{M-m}^{L-\ell} / C_M^L.$$

Найважливішими елементами комбінаторики є розміщення, перестановки і сполучення.

Впорядковані множини даної підмножини (розміщення). Множина A називається ***впорядкованою,*** якщо кожному елементу цієї множини поставлено у відповідність певне число (номер елемента) від 1 до n, де n число елементів множини, так що різним елементам відповідають різні числа.

Теорема 2. ***Множину, яка містить n елементів, можна впорядкувати $P_n = n!$ способами.***

Доведення. Нехай є множина *A,* яка містить *n* елементів. Будемо послідовно вибирати елементи множини *A* і розташовувати їх у певному порядку на *n* місцях. На перше місце можна поставити будь-який з *n* елементів. Після того, як заповнено перше місце, на друге місце можна поставити будь-який з *(n-1)* елементів, які залишились, і так далі. Отже, за правилом множення всі *n* місць можна заповнити способами $n(n-1)\cdot(n-2)\cdot\ldots\cdot 2\cdot 1$. Тому множину *A,* яка містить *n* елементів можна впорядкувати *n!* способами.

Різні впорядкування певної множини називають ***перестановками*** цієї множини. Нехай, наприклад, $A=\{1,2,3\}$. Тоді всі перестановки мають вигляд:

$\{1,2,3\}, \{1,3,2\}, \{2,3,1\}, \{2,1,3\}, \{3,1,2\}, \{3,2,1\}$.

Приклад 9. Скількома способами можна розташувати на полиці 6 книг з різних розділів вищої математики?

Розв'язування. Шукане число способів дорівнює числу способів упорядкування множини, що містить 6 елементів, тобто $P_6 = 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6 = 720$.

Нехай маємо деяку множину A, яка складається з n різних елементів. ***Розміщеннями із n елементів по k елементів*** $(k \le n)$ називаються всі упорядковані підмножини A, які складаються з k різних елементів. Розміщення відрізняються одне від одного або самими елементами, або порядком їхнього розташування. Нехай, наприклад, $A=\{1,2,3,4\}, k=3;$ тоді всі можливі розміщення мають вигляд:

$\{1,2,3\}, \{1,2,4\}, \{1,3,2\}, \{1,3,4\}, \{1,4,2\}, \{1,4,3\},$
$\{2,1,3\}, \{2,3,4\}, \{2,3,1\}, \{2,3,4\}, \{2,4,1\}, \{2,4,3\},$
$\{3,1,2\}, \{3,1,4\}, \{3,2,1\}, \{3,2,4\}, \{3,4,1\}, \{3,4,2\},$
$\{4,1,2\}, \{4,1,3\}, \{4,2,1\}, \{4,2,3\}, \{4,3,1\}, \{4,3,2\}.$

Теорема 3. ***Число розміщень з n елементів по k елементів дорівнює*** $A_n^k = n(n-1)(n-2)\cdot\ldots\cdot(n-k+1).$

Доведення. Утворимо усі можливі розміщення з n елементів по k елементів. При цьому вибір першого елемента можна здійснити n способами; після того як обраний n способами перший елемент, другий елемент можна вибрати із тих, які залишились, (n-1)-м способом, третій елемент - (n-2)- ма способами і т.д.; останній, k-й елемент можна вибрати $n - (k - 1) = n - k + 1$ способами. Так будуть утворені всі можливі розміщення. Згідно основного правила комбінаторики, шукане число розміщень дорівнює $n\cdot(n-1)\cdot\ldots\cdot(n-k+1).$

Розміщення застосовуються при описанні тих об'єктів, для яких важливий порядок розташування елементів.

Приклад 10. Абонент забув 3 останні цифри номера банківської скриньки, але пам'ятає, що усі вони різні. Яка імовірність того, що, набираючи номер навмання, він набере його правильно?

Розв'язування. Очевидно, у номері скриньки відіграє роль порядок цифр. Загальне число способів набрати 3 різні цифри дорівнює $A_{10}^3 = 10 \cdot 9 \cdot 8 = 720$. Нехай подія B={номер набраний правильно}. Очевидно, дана подія B включає одну елементарну подію, отже, $P(B) = \frac{1}{720} \approx 0{,}00139.$.

Розміщення з повтореннями

Нехай множина A складається з n різних елементів. ***Розміщеннями з повтореннями*** із n елементів по k елементів називаються упорядковані підмножини з A, які містять k елементів, в яких елементи можуть повторюватись. Нехай, наприклад, $A = \{ 1,2,3 \}, k = 2$. Тоді всі розміщення з повтореннями мають вигляд:

{1,1}, {1,2}, {2,1}, {2,2}, {1,3}, {3,1}, {2,3}, {3,2}, {3,3}.

Теорема 4. ***Число розміщень з повтореннями із n елементів по k елементів дорівнює $\overline{A}_n^k = n^k$.***

Доведення. Утворимо усі можливі розміщення з повтореннями із n елементів по k елементів. При цьому перший елемент розміщення можна вибрати n способами, другий, в силу можливих повторень елементів, - знову n способами, і так далі. Останній, k-й елемент можна також вибрати n способами. Далі, згідно основного правила комбінаторики, загальне число розміщень з повтореннями дорівнює $\overline{A}_n^k = \underbrace{n \cdot n \cdot \ldots \cdot n}_{k} = n^k$.

Приклад 11. Друкарська машинка має 48 клавіш. Яка ймовірність того, що друкуючи наосліп, можна надрукувати слово "ймовірність"?

Розв'язування. У даній задачі $n = 48,\ k = 11$ (оскільки у слові "ймовірність"- 11 літер). При друкуванні наосліп літери можуть повторюватися, тому загальне число можливих способів надрукувати слово, яке містить 11 літер, дорівнює $n = \overline{A}_{48}^{11} = 48^{11}$. Нехай подія A={ надруковано слово "ймовірність"}. Тоді $|A| = m = 1$, і отже шукана ймовірність дорівнює $P(A) = \frac{1}{48^{11}}$.

Приклад 12. Абонент, набираючи номер телефону, забув дві останні цифри і набрав їх навмання. Яка ймовірність події $A = \{$набрані абонентом навмання дві останні цифри номера телефону є вірні$\}$?

Розв'язування. Число всіх можливих елементарних наслідків випробування є число розміщень з 10 – ти елементів по 2 елементи, тобто

$$n = A_{10}^2 = \frac{10!}{(10-2)!} = \frac{10!}{8!} = \frac{8! \cdot 9 \cdot 10}{8!} = 90.$$

Число сприятливих для події A наслідків $m - 1$. Тому

$$P(A) = \frac{1}{90}.$$

Приклад 13. У відділенні комерційного банку працює 12 економістів і 7 7 бухгалтерів. Керівник відділення вирішив сформувати групу аналізу ризиків з шести осіб: чотирьох економістів і двох бухгалтерів. Яка ймовірність події $A=\{$вибрана навмання група з шести осіб включає чотири економісти і два бухгалтера$\}$?

Розв'язування. Число всіх можливих елементарних наслідків випробування є число комбінацій з 19 – ти осіб по 6 осіб, тобто

$$n=C_{19}^{6}=\frac{19!}{6!(19-6)!}=\frac{15!}{6!\cdot 13!}=\frac{13!\cdot 14\cdot 15\cdot 16\cdot 17\cdot 18\cdot 19}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 13!}=27132\,.$$

Число сприятливих для події A наслідків визначається як добуток числа комбінацій з 12 – ти по 4 економісти і числа комбінацій з 7 – ми по 2 бухгалтера, тобто

$$m=C_{12}^{4}\cdot C_{7}^{2}=\frac{12!}{4!\cdot 8!}\cdot\frac{7!}{2!\cdot 5!}=\frac{8!\cdot 9\cdot 10\cdot 11\cdot 12}{4!\cdot 8!}\cdot\frac{5!\cdot 6\cdot 7}{2!\cdot 5!}=\frac{9\cdot 10\cdot 11\cdot 12}{1\cdot 2\cdot 3\cdot 4}\cdot\frac{6\cdot 7}{1\cdot 2}=10395\,.$$

Отже,

$$P(A)=\frac{m}{n}=\frac{10395}{27132}\approx 0{,}38\ .$$

§5. *Статистичне визначення ймовірності події*

Класичне визначення ймовірності події передбачає, що число елементарних наслідків є скінченне і наслідки є рівноможливі.

На практиці зустрічаються випробування з нескінченним числом можливих наслідків. Крім того, не завжди можливо результат випробування подати у вигляді суми рівноможливих елементарних наслідків. Тобто застосування класичного визначення ймовірності в цих випадках є неможливе.

Наведемо інше визначення ймовірності, яке часто більш зручне для застосування.

Припустимо, що проводиться n однотипних випробувань, одним з наслідків яких є дана подія A.

Означення. ***Відношення числа появи m події A до загального числа випробувань n називається відносною частотою цієї події і позначається $W_n(A)$, тобто***

$$W_n(A)=\frac{m}{n}. \tag{2}$$

Очевидно, $0\le W_n(A)\le 1$.

Спільність між відносною частотою події A та її класичною ймовірністю полягає в тому, що відносна частота є завжди числом, яке близьке до ймовірності події A. Різниця між відносною частотою події A та її класичною ймовірністю полягає в тому, що визначення відносної частоти передбачає проведення випробувань, а визначення класичної ймовірності їх проведення не вимагає.

З формули (2) одержуємо, що $m=W_n\cdot n$, тобто, число появ події A дорівнює його відносній частоті, помноженій на число випробувань.

При однотипних масових випробуваннях спостерігається стійкість відносної частоти події, тобто, якщо число випробувань $n \to \infty$, то відносна частота $W_n(A)$ події A коливається в околі деякого числа p, причому і відхилення від цього числа тим менші, чим більше виконано випробувань. Це число p називають ймовірністю події A в статистичному сенсі.

Означення. ***Статистичною ймовірністю події*** A ***називається границя відносної частоти при необмеженому збільшенні числа випробувань.*** Позначається $W(A)$, тобто

$$W(A) = \lim_{n\to\infty} W_n(A)$$

Отже, цілком достовірно, що відносна частота події наближено співпадає з її статистичною ймовірністю, тому часто за статистичну ймовірність події приймають її відносну частоту.

При широких обмеженнях доводиться, що ймовірності події в класичному і статистичному розуміннях співпадають між собою.

Приклад 14. Аналіз статистики відносної частоти народження хлопчиків певного регіону характеризується такими даними:

січень: 0,514	травень: 0,522	вересень: 0,515
лютий: 0,511	червень: 0,518	жовтень: 0,509
березень: 0,510	липень: 0,538	листопад: 0,518
квітень: 0,529	серпень: 0,516	грудень: 0,521

Середнє арифметичне значення цих частот складає 0,518 і його приймають за ймовірність народження хлопчика.

Приклад 15. При опитуванні тисячі осіб в результаті екзит-полу компанією “Соціальним моніторингом” було визначено, що за кандидата *S* проголосували 428 осіб виборців, за кандидата *T* - 50 виборців, а решта виборців не визначилися. Скільки в середньому виборців проголосували, за кандидата *T*, якщо весь електорат складає 20 млн. осіб, а участь у виборах приймуть 70% виборців?

Розв’язування. За статистичну ймовірність того, що виборець проголосує за кандидата *T*, можна прийняти число $P(\mathrm{T}) = 0{,}05$

У виборах приймуть участь $20 \cdot 0{,}7 = 14$ млн. виборців. Кількість виборців, які проголосують за кандидата *T*, складає $14 \cdot 0{,}05 = 0{,}7$ млн. виборців.

Зауваження. Для існування статистичної ймовірності вимагається виконання двох умов:

а) можливість, хоча б принципово, виконувати нескінченне число випробувань, в кожному з яких подія *A* наступить або не наступить;

б) стійкість відносних частот появи події *A* в різних серіях досить великого числа випробувань, тобто існування границі $\lim_{n\to\infty} W_n(A)$.

§6. *Геометричне визначення ймовірності події*

У випадках, коли класичне визначення ймовірності події не можна застосувати до випробування з нескінченним числом рівноможливих наслідків, вводять поняття геометричної ймовірності.

Нехай G - відрізок числової прямої, або область на площині, або область в просторі, а $g \subset G$ - деяка частина області G. Через $mesG$ і $mesg$ позначимо міру області G і g відповідно (довжина відрізка, площа або об'єм області на площині або в просторі). Нехай подія $A=\{$попадання точки в область g при киданні її навмання в область $G\}$. При цьому припустимо, що ймовірність попадання точки в область пропорційна її мірі і не залежить від її розташування.

Означення. ***Геометричною ймовірністю події*** A називається відношення міри області g до міри області G, тобто

$$P(A)=\frac{mesg}{mesG}. \tag{3}$$

Приклад 16. На відрізок OM довжини L числової осі Ox навмання поставлена точка $B(x)$. Знайти ймовірність того, що менший з відрізків OB і BM має довжину більшу, ніж $\frac{1}{3}L$.

Розв'язування. Нехай подія $A=\{$менший з відрізків OB і BM має довжину більшу ніж $\frac{1}{3}L$,якщо точку B навмання поставити на відрізок $OM\}$.

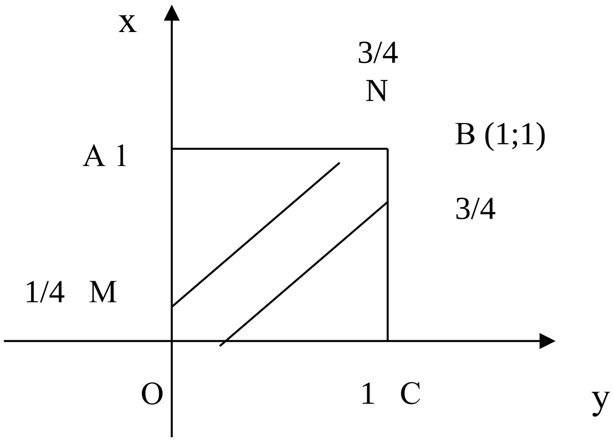
Подія A відбудеться, якщо точка B попаде на відрізок $CD \subset OM$, кінці якого C і D ділять відрізок OM на 3 рівні частини. Отже шукана ймовірність

$$P(A)=\frac{mesCD}{mesOM}=\frac{\frac{1}{3}L}{L}=\frac{1}{3}.$$

Приклад 17. (***класична задача про зустріч***) Два студенти домовилися про зустріч у визначеному місці між 12 і 13 годинами. Студент, який приходить на зустріч першим, чекає свого колегу лише 15 хв. Знайти ймовірність того, що студенти зустрінуться у визначеному місці, якщо кожний студент вибирає навмання час приходу на місце зустрічі між 12 і 13 годинами.

Розв'язування. Нехай подія $S=\{$студенти зустрінуться у визначеному місці у випадку виконання описаних в задачі умов зустрічі$\}$. Позначимо через $12+x$ - час приходу першого студента, $12+y$ - час приходу другого студента на місце зустрічі після 12 год. Очевидно, $0 \le x \le 1, \quad 0 \le y \le 1$.

Розглянемо систему координат Oxy і побудуємо квадрат $0 \le x,y \le 1$ (мал.2). Зустріч студентів відбудеться, якщо $(x,y) \in S$



Мал. 2.

Проводимо прямі $y = x + \frac{1}{4}$ і $y = x - \frac{1}{4}$ і приходимо до висновку: студенти зустрінуться у визначеному місці, коли точка (x, y) попаде у заштриховану область (див. рис. 2). Тому,

$$P(S) = \frac{mesOMNBQP}{mesOABC} = \frac{\textit{пл.кв.}\ \text{OABC} - 2\ \text{пл.}\ \Delta\text{MAN}}{\text{пл.кв.OABC}} =$$

$$= \frac{1 - 2 \cdot \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{3}{4}}{1} = 1 - \frac{9}{16} = \frac{7}{16} = 0{,}4375\,.$$

Зауваження. При розв’язуванні багатьох практичних задач ми зустрічаємося з подіями, ймовірність яких дуже мала. Чи можна вважати, що така подія A в одиничному випробуванні не відбудеться? Такого висновку зробити не можна, бо не виключено, що подія A відбудеться.

Як показує практичний досвід, малоймовірні події в одиничному випробуванні в переважній більшості не відбуваються. На основі цього факту приймають такий ***“принцип практичної неможливості малоймовірних подій”: якщо випадкова подія має дуже малу ймовірність, то практично можна вважати, що в одиничному випробуванні ця подія не відбудеться.***

Наскільки малою повинна бути ймовірність події, щоб можна було вважати її появу в одиничному випробуванні неможливою, залежить від конкретної задачі. Якщо, наприклад, ймовірність неповернення позики дорівнює *0,1*, то було б недопустимо вважати, що така подія не відбудеться; якщо ж ймовірність неповернення позики *0,01*, то можна вважати, що вона буде сплачена вчасно.

Досить мала ймовірність, при якій (в даній конкретній задачі) подію можна вважати неможливою, називається ***рівнем значущості***, який на практиці вважають числом, що міститься між *0,01* і *0,05*.

Аналогічно можна сформулювати висновок про події, ймовірність яких є близькою до одиниці: ***якщо випадкова подія має ймовірність близьку до***

одиниці, то можна вважати, що в одиничному випробуванні ця подія відбудеться.

Завдання для самостійної роботи студентів

1. Гральний кубик кидають двічі. Описати простір елементарних подій. Описати наступні події: $A=${принаймні один раз з'явиться цифра шість}; $B=${при другому підкиданні з'явиться цифра шість}; $A\cup B$; $B\cap A$; $B\setminus A$; $\overline{B}$.

2. Розглянемо події: $A=${взята навмання деталь є першого сорту}, $B=${ взята навмання деталь є другого сорту }, $C=${ взята навмання деталь є третього сорту }. Описати словами події: 1) $A\cup B$; 2) $\overline{A\cup B}$; 3) $A\cup C$; 4) $A\cup B\cup C$.

3. Рада директорів компанії складається з трьох економістів, двох менеджерів і двох маркетологів. У нову раду повинно ввійти три особи. Знайдіть імовірність того, що навмання обрана рада буде складатися:
а) повністю з економістів; б) менеджера та двох економістів.

4. Обчислити ймовірність того, що дні народження 12 осіб припадуть на різні місяці року (число сприятливих наслідків – перестановки, а число всіх наслідків – розміщення з повторенням).

5. В партії з 500 деталей відділ технічного контролю виявив 9 нестандартних деталей. Знайти відносну частоту появи нестандартної деталі. Скільки можна виявити нестандартних деталей у партії з 1000 деталей?

6. Яка ймовірність виграти в лотерею 6 із 39?, 5 із 36?

7. В результаті ряду спроб було визначено, що при 250 пострілах стрілець попадає в ціль в середньому 230 разів. Яка ймовірність враження цілі цим стрільцем? Скільки можна очікувати від нього попадань в ціль при 1000 пострілах?

8. Таємний код банківської картки складається з чотирьох цифр. Ви забули свій код, але пам'ятаєте, що дві цифри однакові. Яка ймовірність того, що навмання набраний Вами код є вірним?

9. Навмання взятий телефонний номер складається з семи цифр. Яка ймовірність того, що в цьому номері кожна цифра ділиться на три?

10. В результаті соціологічного опитування 3000 осіб було визначено, що за кандидата *X* проголосує 946 виборців, за кандидата *У* – 1054, решта виборців ще не визначилась. Знайти ймовірність вибору кожного з кандидатів. Скільки всередньому осіб проголосує за кандидата *X*, якщо весь електорат складає 12 млн. чол. при 72% активності виборців.

11. В коло вписаний квадрат. В круг, обмежений цим колом, кидають навмання точку. Яка ймовірність того, що вона попаде в квадрат?

12. В рівносторонній трикутник вписано круг. В цей трикутник кидається навмання точка. Яка ймовірність того, що вона не попаде в круг?

13. Для розвантаження два сухогрузи повинні підійти до одного приплаву. Їх поява – незалежні випадкові події, рівноможливі протягом доби.

Знайти ймовірність того, що одному із сухогрузів доведеться чекати звільнення приплаву, якщо час стоянки першого одна година, а другого – дві години.

14. Сім академічних груп студентів розмістили в семи аудиторіях, що знаходяться поруч. Знайдіть імовірність того, що перша та друга групи знаходяться в сусідніх аудиторіях.

15. Кожного четверга інкасаторська машина доставляє гроші з відділення банку в чотири фірми. Для убезпечення перевезень використовують різні маршрути. Водій навмання обирає один із варіантів. Знайдіть імовірність того, що: а) нинішній маршрут не повторює попередній; б) після фірми А гроші буде відразу доставлена в фірму В.

16. З колоди гральних карт, яка складається з 36 карт, навмання взяли чотири карти. Знайдіть імовірність того, що: а). серед взятих карт усі тузи; б). рівно один туз; в) хоча б один туз; г). всі карти однієї масті.

17. Знайдіть імовірність того, що при киданні двох гральних кубиків: а) сума чисел, що випали, буде дорівнювати п'яти, б) першою випала одиниця, в) добуток чисел, що випали, дорівнює шести.

18. При перевірці 400 деталей виявлено, що відносна частота бракованих деталей дорівнює 0,07. Скільки виявлено бракованих деталей?

19. Відділ технічного контролю типографії виявив 6 бракованих книг в партії випадково відібраних 120 книг. Знайдіть відносну частоту появи бракованих книг.

20. Серед 50 деталей є 7 бракованих. Навмання вибирають 5 деталей. Знайдіть ймовірність того, що серед них рівно 2 браковані.

21. Автобуси певного маршруту ходять за розкладом. Інтервал руху 15 хв. Знайдіть ймовірність того, що пасажир, який підійшов до зупинки, чекатиме наступний автобус менше 4 хв.

22. На площину з паралельними прямими, відстань між якими $2a$, навмання кинута монета радіуса $r < a$. Яка ймовірність того, що монета не перетне жодної з прямих?

Глава 2. Основні теореми теорії ймовірностей

§1. Теорема додавання ймовірностей

В §1. глави 1 визначено, що сумою двох подій A і B називають подію $A \cup B$, яка полягає в тому, що принаймні одна з подій A або B відбудеться, тобто відбудеться або подія A, або подія B, або обидві події A і B.

Наприклад, якщо покупець придбав два вироби і подія $A=${перший виріб якісний}, подія $B=${другий виріб якісний}, то подія $A \cup B=${перший виріб якісний, або другий виріб якісний, або обидва вироби якісні}.

Теорема 5. ***Ймовірність суми двох несумісних подій дорівнює сумі ймовірностей цих подій, тобто, якщо*** $A \cap B = \emptyset$***, то***

$$P(A \cup B) = P(A) + P(B). \tag{4}$$

Доведення. Нехай в числі n рівноможливих елементарних наслідків випробування сприятливими для події A є m_1 і для події B - m_2 наслідків. Оскільки події A і B несумісні, тобто поява однієї виключає появу другої, то число сприятливих наслідків для події $A \cup B$ дорівнює $m_1 + m_2$ за формулою (1) маємо:

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

Наслідок. Ймовірність суми скінченного числа попарно несумісних подій дорівнює сумі ймовірностей цих подій, якщо події $A_1, A_2, \ldots, A_n$ попарно несумісні, то

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

Обгрунтуємо наслідок для випадку трьох попарно несумісних подій A, B, C. Оскільки події A, B, C попарно несумісні, тобто $A \cap B = \emptyset$, $A \cap C = \emptyset$, $C \cap B = \emptyset$, то події $A \cup B$ і C також несумісні, оскільки $(A \cup B) \cap C = (A \cap C) \cup (B \cap C) = \emptyset$.

Далі за доведеною теоремою маємо:

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

Приклад 18. „Промінвестбанк" оцінює стан кредитування клієнта, що визначає рейтинг, як "відмінний", "добрий", "задовільний" і "незадовільний". Ймовірність того, що клієнт отримає "відмінний" рейтинг дорівнює 0,2, "добрий" – 0,25, "задовільний" – 0,3.

Яка ймовірність того, що:

а) клієнт отримає " незадовільний" рейтинг;

б) клієнт отримає не більше, ніж "добрий" рейтинг?

Розв'язування. Визначимо події:

$A =$ { клієнт отримала "відмінний" рейтинг};

$B =$ { клієнт отримала "добрий" рейтинг};

$C =$ { клієнт отримала "задовільний" рейтинг};

$D =$ { клієнт отримала "незадовільний" рейтинг};

$E=\{$ клієнт отримала не більше, ніж “добрий” рейтинг$\}$.

А). Очевидно, подія $A\cup B\cup C\cup D$ є достовірна і $P(A\cup B\cup C\cup D)=1$. Враховуючи, що події A, B, C, D несумісні, за теоремою про ймовірність суми маємо:

$$P(A\cup B\cup C\cup D)=P(A)+P(B)+P(C)+P(D)=1,$$

а отже,

$$P(D)=1-P(A)-P(B)-P(C)=1-0{,}2-0{,}35-0{,}3=0{,}15.$$

Тобто ймовірність того, що клієнт отримає “незадовільний” рейтинг, дорівнює 0,15.

Б). Подія $E=D\cup C\cup B$ і $P(E)=P(D\cup C\cup B)=P(D)+P(C)+P(B)=$ $=0{,}15+0{,}3+0{,}35=0{,}8$, тобто ймовірність клієнта отримати не більше ніж “добрий” рейтинг дорівнює 0,8.

§2. *Повна група подій*

Розглянемо довільну сукупність подій $A_1, A_2, \ldots, A_n$.

Означення. ***Система подій $A_1, A_2, \ldots, A_n$ називається повною групою подій для даного випробування, якщо будь – яким наслідком є одна і тільки одна подія групи.***

Іншими словами, для повної групи подій виконуються наступні умови:

а) подія $A_1\cup A_2\cup\ldots\cup A_n$ достовірна;

б) події A_i і A_j $(i\neq j)$ попарно несумісні, тобто $A_i\cap A_j=\emptyset$, $\emptyset$ - неможлива подія.

Теорема 6. ***Сума ймовірностей повної групи подій дорівнює одиниці.***

Доведення. Для повної групи подій $A_1, A_2, \ldots, A_n$ подія $A_1\cup A_2\cup\ldots\cup A_n$ достовірна і тому

$$P(A_1\cup A_2\cup\ldots\cup A_n)=1.$$

Оскільки події повної групи попарно несумісні, то

$$P(A_1\cup A_2\cup\ldots\cup A_n)=P(A_1)+P(A_2)+\ldots+P(A_n)$$

Звідси маємо:

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

Теорема доведена.

Приклад 19. Центр тестування школярів отримує тестові роботи з трьох міст M_1, M_2, M_3. Ймовірність отримання листа з міста M_1 складає $p_1=0{,}40$, з міста M_2 - $p_2=0{,}50$. Яка ймовірність того, що наступний лист надійде з третього міста M_3.

Розв’язування. Розглянемо події: $A_1=\{$лист надійшов з міста $M_1\}$, $A_2=\{$ лист надійшов з міста $M_2\}$, $A_3=\{$ лист надійшов з міста $M_3\}$. Події A_1, A_2, A_3 утворюють повну групу. Отже

$$P(A_1)+P(A_2)+P(A_3)=1,$$

Шукана ймовірність буде,

$$P(A_3) = 1 - P(A_1) - P(A_2) = 1 - 0{,}40 - 0{,}50 = 0{,}10.$$

§3. *Ймовірність протилежної події*

Подія A та протилежна подія $\overline{A} = \Omega \setminus A$ утворюють повну групу, оскільки $A + \overline{A} = \Omega$, а простір Ω елементарних подій – повна група. Звідси маємо, що

$$P(A) + P(\overline{A}) = 1 \Leftrightarrow P(\overline{A}) = 1 - P(A),$$

тобто ***сума ймовірностей протилежних подій дорівнює одиниці.***

Приклад 20. Кредитний менеджер комерційного банку аналізує 50 заявок, з яких 30 від малих підприємств. Знайти ймовірність того, що серед чотирьох навмання взятих заявок хоча б одна з малого підприємства.

Розв'язування. Події $A = \{$серед чотирьох навмання взятих заявок хоча б одна з малого підприємства$\}$ і $\overline{A} = \{$всі з чотирьох заявок поступили не з малого підприємства$\}$ є протилежними. Отже, $P(A) = 1 - P(\overline{A})$. Обчислимо ймовірність $P(\overline{A})$. Кількість всіх наслідків випробування є число всіх способів C_{50}^4, якими можна з 60 заявок обрати 4. Кількість сприятливих для події $P(\overline{A})$ наслідків є число способів C_{30}^4, якими можна обрати 4 із 30 заявок від малих підприємств. Тому

$$P(\overline{A}) = \frac{C_{30}^4}{C_{50}^4} = \frac{30!}{4! \cdot 26!} \cdot \frac{4! \cdot 46!}{50!} = \frac{\cdot 27 \cdot 28 \cdot 29 \cdot 30}{1} \cdot \frac{1}{\cdot 47 \cdot 48 \cdot 49 \cdot 50} \approx 0{,}0369.$$

Тому шукана ймовірність

$$P(A) = 1 - 0{,}0369 = 0{,}9631.$$

§4. *Умовна ймовірність події*

Нехай випадковий експеримент проводиться у два етапи. На першому етапі може відбутися або не відбутися деяка подія A, і настання або не настання події A впливає на другий етап експерименту. Тоді імовірність деякої події B, яка може відбутися або не відбутися на другому етапі експерименту, обчислена за умови, що подія A відбулася (умовна імовірність події B за умови A, позначається $P_A(B)$) або $P(B/A)$ і буде відрізнятися від звичайної (безумовної) імовірності події $B, P(B)$.

Означення. Нехай імовірність події A не дорівнює нулю. ***Умовною ймовірністю події*** B ***за умови, що подія*** A ***відбулася, називається число***

$$P(B/A) = \frac{P(A \cap B)}{P(A)}. \qquad (5)$$

Якщо $P(A) = 0$, то умовна імовірність $P(B/A)$ не визначена. В цьому випадку її покладають рівною нулю.

Приклад 21. У групі КІБСа навчається 30 студентів. З них 9 студентів вивчають лише бухоблік, 6 студентів - лише економіку, інші – бухоблік і економіку. Яка імовірність того, що:

а) студент, обраний навмання, вивчає бухоблік?

б) студент, обраний навмання, вивчає економіку?

в) студент, обраний навмання, вивчає економіку. Яка імовірність того, що він вивчає також і бухоблік?

Розв'язування. а) Розглянемо подію $B=${студент, обраний навмання, вивчає бухоблік}. Оскільки із загального числа 30 студентів 24 вивчають бухоблік, то $P(B)=\frac{24}{30}=0{,}8.$

б) Нехай подія $A=${студент, обраний навмання, вивчає економіку}. Очевидно, що $P(A)=\frac{21}{30}=0{,}7.$

в) Далі випадковий експеримент проводиться у два етапи: спочатку з'ясовується знання економіка, потім – бухобліку. На першому етапі відбувається подія A - студент, обраний навмання, вивчає економіку. Імовірність події B обчислюється за умови, що відбулася подія A. Оскільки студентів, які вивчають економіку – 21, з них 15 вивчають також і бухоблік, то $P(B/A)=\frac{15}{21}\approx 0{,}71.$

Зауважимо, що у даному випадку $P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{15}{30}:\frac{21}{30}=\frac{15}{21}$.

Приклад 22. В регіоні є 30 фірм і тільки 10 з них користуються кредитами ПІБу. Із списку фірм навмання вибирають дві підряд. Знайти ймовірність того, що:

а) друга фірма не користується кредитом банку (подія A), коли перша фірма ним користується (подія B).

б) друга фірма користується кредитом банку (подія A), коли перша фірма також ним користується (подія B).

Розв'язування. Зазначимо, що обидві шукані ймовірності є умовними:

а) якщо першою вибрана фірма, яка користується кредитом (подія B відбулася), то $P\left(A/B\right)=\frac{20}{29}\approx 0{,}6896$;

б) якщо першою вибрана фірма, яка не користується кредитом (подія B не відбулася), то $P\left(A/\overline{B}\right)=\frac{9}{29}\approx 0{,}3103$.

Властивості умовної ймовірності події

Безпосередньо із означення, маємо:

1. $0\le P(B/A)\le 1,\ P(A/A)=1.$

2. Якщо події B і C несумісні, то

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

Із означення умовної імовірності слідує ***формула множення ймовірностей для будь-яких двох подій*** A ***і*** B ***з одного і того ж простору елементарних подій:***

Теорема 7. ***Ймовірність добутку подій*** A ***і*** B ***дорівнює добутку ймовірності однієї з них на умовну ймовірність другої, за умови, що перша подія відбулася:***

$$P(A\cap B)=P(B/A)\cdot P(A)=P(A/B)\cdot P(B) \tag{6}$$

Доведення. Проведемо доведення теореми згідно класичної схеми. Нехай для події A сприятливими є m, а для події AB - k рівноможливих елементарних наслідків із загального їх числа n. Тоді

$$P(A)=\frac{m}{n} \quad \text{і} \quad P(A\cap B)=\frac{k}{n}$$

Якщо подія A відбулася, то в цій ситуації можливі лише ті m елементарних наслідків, які були сприятливими для події A, причому k з них є сприятливі для події B. Звідси маємо, що

$$P\,(B/A)=\frac{k}{m}.$$

Використовуючи ці три рівності, одержуємо:

$$P(A\cap B)=\frac{k}{n}=\frac{m}{n}\cdot\frac{k}{m}=P(A)P\,(B/A).$$

Оскільки $A\cap B=B\cap A$, то

$$P(A\cap B)=P(B\cap A)=P(B)P\,(A/B).$$

Теорема доведена.

Означення. ***Дві події*** A ***і*** B ***називаються незалежними, якщо ймовірність кожної з них не залежить від появи або не появи іншої, тобто***

$$P(A)=P\,(A/B)=P\,(A/\overline{B}) \ \ \textbf{і} \ \ P(B)=P\,(B/A)=P(B/\overline{A}).$$

В протилежному випадку події називаються ***залежними.***

Наслідок. ***Ймовірність добутку двох незалежних подій*** A ***і*** B ***дорівнює добутку ймовірностей цих подій***:

$$P(A\cap B)=P(A)\cdot P(B). \tag{7}$$

Формула (7) випливає безпосередньо з формули (6), оскільки в цьому ипадку $P_A(B/A)=P(B)$ і $P\,(A/B)=P(A)$.

Зауваження. Дану теорему множення ймовірностей можна узагальнити на випадок довільного числа подій $A_1, A_2, \ldots, A_n$ і вона виражається формулою

$$P(A_1\cap A_2\cap\ldots\cap A_n)=P(A_1)\cdot P_1\,(A_2/A_1)\cdot P_2\,(A_3/A_1\cap A_2)\cdot\ldots\cdot P\,(A_n/A_1\cap\ldots\cap A_{n-1})$$

Дана формула спрощується, якщо події $A_1, A_2, \ldots, A_n$ ***незалежні у сукупності,*** тобто, якщо вони попарно незалежні й незалежна кожна подія та всі можливі добутки інших. Наприклад, події A_1, A_2, A_3 незалежні у сукупності, якщо незалежні події A_1 і A_2, A_1 і A_3, A_2 і A_3, A_1 і $A_2\cap A_3$, A_2 і $A_1\cap A_3$, A_3 і $A_1\cap A_2$. Легко переконатися в тому, що незалежні події у сукупності попарно незалежні. Звернемо увагу на те, що обернене твердження невірне.

Аналогічно, як у випадку двох незалежних подій, маємо: ***якщо події*** $A_1, A_2, ..., A_n$ ***незалежні у сукупності,*** то

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1) \cdot P(A_2) \cdot ... \cdot P(A_n). \qquad (8)$$

Нехай $A_1, A_2, ..., A_n$ - ***довільні події***, задані на одному і тому ж просторі елементарних подій Ω. Тоді

$$P(\bigcap_{i=1}^{n} A_i) = P(A_1) \cdot P(A_2 / A_1) \cdot P(A_3 / A_1 \cap A_2) \cdot ... \cdot P(A_n / \bigcap_{i=1}^{n-1} A_i).$$

Дійсно, відповідно до формули (6), маємо

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_n / A_1 \cap A_2 \cap ... \cap A_{n-1}) P(A_1 \cap ... \cap A_{n-1}) =$$
$$= P(A_n / A_1 \cap A_2 \cap ... \cap A_{n-1}) P(A_{n-1} / A_1 \cap A_2 \cap ... \cap A_{n-2}) \times$$
$$\times P(A_1 \cap A_2 \cap ... \cap A_{n-2}) = ... =$$
$$= P(A_n / A_1 \cap ... \cap A_{n-1}) P(A_{n-1} / A_1 \cap ... \cap A_{n-2}) \cdot ... \cdot P(A_1).$$

Приклад 23. Одеське відділення ПІБу отримало 15 заявок від підприємств на надання позик. Серед них 6 заявок – від міських підприємств, і 9 – від обласних. Розглядаються підряд 4 заявки. Яка імовірність того, що усі вони – від обласних підприємств?

Розв'язування. Нехай події A_i $(i = 1,2,3,4)$ полягають у тому, що $i-$а вийнята заявка – від обласних підприємств. Потрібно обчислити ймовірність $P(A_1 \cap A_2 \cap A_3 \cap A_4)$.

Згідно формули (7) запишемо:

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1) P(A_2 / A_1) P(A_3 / A_1 \cap A_2)$$
$$P(A_4 / A_1 \cap A_2 \cap A_3) = \frac{9}{15} \cdot \frac{8}{14} \cdot \frac{7}{13} \cdot \frac{6}{12} = \frac{12}{65}.$$

Приклад 24. Колектив працівників комерційного банку складається на 68% з жінок. Серед працівників 30% чоловіків і 36% жінок мають вищу банківську освіту. Знайти ймовірність того, що навмання вибраний працівник банку є чоловік і має вищу банківську освіту.

Розв'язування. Розглянемо події: $A = \{$вибраний навмання працівник банку має вищу банківську освіту$\}$, $B = \{$вибраний навмання працівник є чоловік$\}$, $A \cap B = \{$вибраний навмання працівник банку є чоловік і має вищу банківську освіту$\}$. За умовою задачі $P(B) = 0,32$; $P(A / B) = 0,3$. За формулою (6) шукана ймовірність події $A \cap B$ дорівнює:

$$P(B \cap A) = P(B) \cdot P(A / B) = 0,32 \cdot 0,3 = 0,105.$$

Приклад 25. Центральне відділення ПІБу аналізує стан погашення кредитів по своїх регіональних відділень. Умовою наскрізної перевірки відділення банку є наявність хоча б однієї проблемної позики (не погашення кредиту) серед п'яти навмання взятих для перевірки. Сто позик комерційного банку піддають вибірковому контролю. Яка ймовірність того, що портфель позик буде проблемним, якщо він містить 3% ризикованих позик.

Розв'язування. Розглянемо події: $A = \{$портфель позик проблемний$\}$, $A_i = \{i$-та перевірена позика не проблемна$\}$, $i = \overline{1,5}$. Подія

$\overline{A} = A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5$ є протилежна до події A. Обчислимо її ймовірність. Згідно з правилом множення, маємо:

$$P(\overline{A}) = P(A_1) \cdot P_1(A_2 / A_1) \cdot P_2(A_3 / A_1 \cap A_2) \cdot P_3(A_4 / A_1 \cap A_2 \cap A_3) \cdot$$
$$\cdot P_4(A_5 / A_1 \cap A_2 \cap A_3 \cap A_4)$$

За умовою задачі в портфелі зі 100 позик 97 не проблемні. Отже, $P(A_1) = \frac{97}{100}$. Після того, коли подія A_1 відбулася (перша вибрана позика не проблемна), і в партії залишилося 99 позик, серед яких 96 не проблемних, то $P_A(A_2 / A_1) = \frac{96}{99}$.

Аналогічно обчислимо:

$$P_2(A_3 / A_1 \cap A_2) = \frac{95}{98}, \quad P_2(A_4 / A_1 \cap A_2 \cap A_3) = \frac{94}{97}$$
$$P_2(A_5 / A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{93}{96}.$$

Шукану ймовірність знайдемо за формулою $P(A) = 1 - P(\overline{A})$:

$$P(A) = 1 - \frac{97}{100} \cdot \frac{96}{99} \cdot \frac{95}{98} \cdot \frac{94}{97} \cdot \frac{93}{96} = 1 - 0{,}87 = 0{,}13.$$

Приклад 26. Стан використання підприємствами наданих їм кредитів щорічно контролюється центральним та обласним банками. Центральний банк перевіряє вибірково 15%, а обласний банк - 21% підприємств. Вважаючи, що рішення щодо перевірки на різних рівнях приймаються незалежно, знайти ймовірність того, що підприємство перевірятиметься в поточному році центральним банком і не перевірятиметься обласним банком.

Розв'язування. Розглянемо події: $A = \{$підприємство перевіряється центральним банком$\}$, $B = \{$підприємство перевіряється обласним банком$\}$. Подія $A \cap \overline{B}$ полягає в тому, що підприємство перевіряється центральним банком і не перевіряється обласним банком.

В силу незалежності подій A і $\overline{B}$, шукану ймовірність обчислюємо за формулою (7):

$$P(A \cdot \overline{B}) = P(A) \cdot P(\overline{B}) = 0{,}15 \cdot (1 - 0{,}21) = 0{,}15 \cdot 0{,}79 = 0{,}1185.$$

§5. *Ймовірність появи принаймні однієї події*

Нехай в результаті випробування можуть з'явитися n подій: $A_1, A_2, ..., A_n$, які незалежні у сукупності, ймовірності появи кожної з цих подій відомі. Як знайти ймовірність появи хоча б однієї з цих подій? Наприклад, якщо в результаті випробування можуть з'явитися три події, то поява хоча б однієї з цих подій означає появу або однієї, або двох, або трьох подій. Відповідь на це питання дає наступна теорема.

Теорема 8. ***Ймовірність появи хоча б однієї з подій $A_1, A_2, \ldots, A_n$, які незалежні у сукупності, дорівнює різниці між одиницею і добутком ймовірностей протилежних подій, тобто ймовірність події $A =$ {поява хоча б однієї з подій $A_1, A_2, ..., A_n$} обчислюється за формулою:***

$$P(A) = 1 - P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot \ldots \cdot P(\overline{A}_n). \tag{9}$$

Доведення. Події A і $B = \overline{A}_1 \cap \overline{A}_2 \cap \ldots \cap \overline{A}_n$ (кожна з подій $A_1, A_2, \ldots, A_n$ не відбудеться) є протилежні і отже:

$$P(A) = 1 - P(B) = 1 - P(\overline{A}_1 \cap \overline{A}_2 \cap \ldots \cap \overline{A}_n).$$

Оскільки події $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$ несумісні у сукупності, то за формулою множення ймовірностей, маємо

$$P(\overline{A}_1 \cap \overline{A}_2 \cap \ldots \cap \overline{A}_n) = P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot \ldots \cdot P(\overline{A}_n).$$

Звідки,

$$P(A) = 1 - P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot \ldots \cdot P(\overline{A}_n).$$

Приклад 27. Підприємець має договорні відносини з трьома банками й може звертатися за кредитом до кожного з них. Протягом попередніх п'яти років перший банк виділяв кредити 6 разів, другий – 8 разів, третій – 9 разів при десяти звертаннях до кожного з них. Яка ймовірність того, що в даний момент хоча б один з банків виділить підприємцеві кредит?

Розв'язування. Позначимо події: $A_i = \{i$-тий банк виділить бізнесменові кредит$\}$, $(i = 1,2,3)$, $A = \{$хоча б один з банків A_1, A_2, A_3 виділить бізнесменові кредит$\}$.

Очевидно, $P(\overline{A}_1) = \dfrac{4}{10}$, $P(\overline{A}_2) = \dfrac{2}{10}$, $P(\overline{A}_3) = \dfrac{1}{10}$. За формулою (9) шукана ймовірність

$$P(A) = 1 - P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot P(\overline{A}_3) = 1 - \frac{4}{10} \cdot \frac{2}{10} \cdot \frac{1}{10} = 1 - 0{,}008 = 0{,}982.$$

§6. *Теорема додавання ймовірностей довільних подій*

Вище ми розглянули теорему додавання ймовірностей несумісних подій, яка виражається формулою (4), §1 даної глави.

Припустимо тепер, що події A і B ***сумісні***, тобто поява однієї з них не виключає появи іншої в одному й тому ж випробуванні. Як знайти ймовірність події $A \cup B$, яка полягає в тому, що відбудеться хоча б одна з подій A і B?

Теорема 9. ***Ймовірність суми двох сумісних подій дорівнює сумі ймовірностей цих подій без ймовірності їх сумісної появи:***

$$P(A \cup B) = P(A) + P(B) - P(AB). \tag{10}$$

Доведення. Оскільки події A і B несумісні, то подія $A \cup B$ відбудеться, коли відбудеться одна з трьох несумісних подій $\overline{A} \cap B$, $A \cap \overline{B}$, $A \cap B$, тобто

$$A \cup B = \overline{A} \cap B + A \cap \overline{B} + A \cap B.$$

За теоремою додавання ймовірностей несумісних подій, маємо

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

Подія A відбудеться, коли відбудеться одна з несумісних подій $A\cap\overline{B}$ і $A\cap B$, або $A=A\cap\overline{B}\cup A\cap B$. Згідно з теоремою додавання ймовірностей несумісних подій

$$P(A)=P\left(A\cap\overline{B}\right)+P(A\cap B)$$

Аналогічно одержимо, що

$$P(B)=P\left(\overline{A}\cap B\right)+P(A\cap B).$$

З попередніх двох рівностей маємо:

$$P\left(A\cap\overline{B}\right)=P(A)-P(A\cap B),\quad P\left(\overline{A}\cap B\right)=P(B)-P(A\cap B)$$

і, це дозволяє записати

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

Теорема доведена.

Зауваження 1. Формула (10) справедлива як для залежних, так і незалежних подій. При цьому, якщо події A і B незалежні, то $P(A\cap B)=P(A)\cdot P(B)$ і формула (10) набуває вигляду:

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

Зауваження 2. З формули (10), як частинний випадок, дістаємо формулу для ймовірності суми несумісних подій, бо в цьому випадку $P(A\cap B)=0$ і формула (10) зводиться до вигляду

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

Приклад 28. Два бізнесмени взяли кредит у банку. Ймовірність того, що перший бізнесмен погасить свій вчасно кредит, дорівнює 0,73, а другий – 0,65. Знайти ймовірність того, що хоча б один з бізнесменів погасить вчасно свій кредит?

Розв'язування. Розглянемо події: $A_1=\{$перший підприємець погасить вчасно свій кредит $\}$, $A_2=\{$другий підприємець погасить вчасно свій кредит $\}$. Тоді подія, яка полягає в тому, що хоча б один підприємець погасить кредит вчасно, є сума сумісних і незалежних подій A_1 і A_2. Шукану ймовірність події $A_1\cup A_2$ обчислюємо за формулою (10):

$$P(A_1\cup A_2)=0{,}73+0{,}65-0{,}73\cdot 0{,}65=0{,}9055.$$

Завдання для самостійної роботи студентів

1. На студента чекають іспити з англійської мови та математики. На його думку ймовірність того, що він складе іспит з англійської мови, дорівнює 0,6; ймовірність того, що він здасть принаймні один предмет, дорівнює 0,9, а ймовірність того, що здасть обидва предмети – лише 0,5. Яка ймовірність того, що він складе іспит з математики?

2. Власник кафе звернув увагу, що 75 % відвідувачів беруть після обіду десерт, 60 % – соки, 40 % – те й інше. Яка ймовірність, що навмання обраний відвідувач, який замовив десерт, візьме сік?

3. Чотири фірми беруть участь у конкурсі на отримання контракту. Ймовірність отримати контракт для кожної з них відповідно дорівнюють 0,2;

0,4; 0,3 і 0,1. Стало відомо, що перша фірма брати участь в конкурсі не буде. Як змінилися шанси інших трьох фірм?

4. Є два ринки цінних паперів. Фірма направила на обидва ринки *260* акцій водночас різноманітної якості. У визначений день на ринку 1 зросли ціни *197* акцій (подія А), а на ринку 2 зросли ціни *191* акції (подія В). Одночасно на обох ринках зросли ціни *165* акцій. Інші акції не піднялися в ціні. Результат торгів наведений у таблиці:

	A	$\overline{A}$	Разом
B	*165*		*191*
$\overline{B}$			
Разом	*197*		*260*

Заповніть до кінця таблицю. Визначіть імовірності таких подій: а) ціна акцій не зросла одночасно на обох ринках, б) ціна акцій зросла на ринку 2, в) ціна акцій зросла на ринку 1 або 2, г) зростання акцій на другому ринку, якщо відомо, що на першому ринку їх ціна не піднялася, д) зростання акцій на першому ринку, якщо відомо, що їх ціна зросла на другому ринку.

5. З опитаних бізнесменів 80 % віддають перевагу зберіганню грошей у банку (подія В), 60 % вкладають гроші в цінні папери (подія С), 50 % одночасно тримають гроші в банку та у вигляді цінних паперів. Результати опитування наведені в таблиці:

	B	$\overline{B}$	Разом
C	*0,5*		*0,6*
$\overline{C}$			
Разом	*0,8*		*1*

Слід до кінця заповнити таблицю. Визначіть імовірність того, що навмання обраний бізнесмен: а) тримає гроші в банку або у вигляді цінних паперів; б) не тримає гроші в банку за умови, що він не тримає цінних паперів.

6. На іспиті з економіки студент оцінює ймовірність отримання оцінки “відмінно” в *0,25* і ймовірність отримання оцінки “добре” – *0,36*. Яка ймовірність, що студент:

1). не отримає оцінки “відмінно”?

2). не отримає оцінки “добре”?

3). не отримає ні “відмінно”, ні “добре”?

7. Ймовірність погашення кредиту для першого позичальника становить *0,75*, а для другого – *0,85*. Визначити ймовірність своєчасного погашення кредиту хоча б одним підприємцем, якщо вони працюють незалежно.

8. Ймовірність банкрутства для першого підриємства дорівнює *0,24*, а для другого ця ймовірність на *15%* більша. Знайти ймовірність того, що із двох підприємств збанкрутує хоча б одне.

9. Ймовірність безвідмовної роботи блока, що входить в систему, протягом певного часу дорівнює *0,75*. Для підвищення надійності встановлюють такий самий резервний блок. Знайти ймовірність безвідмовної роботи системи з врахуванням резервного блоку.

10.На ділянці дороги *AB* для водій спортивної машини має сім перешкод. Ймовірність зупинки на кожній з них дорівнює *0,15*, а ймовірність того, що водій проїде з пункту *B* до кінцевого пункту *C* без зупинки, становить *0,65*. Обчислити ймовірність того, що на ділянці дороги *AC* не буде зупинок.

11.Відділом технічного контролю при прийомі партії виробів перевіряється третина виробів. Умовами прийому допускається не більше *3%* бракованих виробів. Визначити ймовірність того, що партія з *600* виробів буде прийнята, якщо вона містить *6%* браку.

12.Два пункти сполучаються кількома лініями зв'язку і ймовірність пошкодження кожної з них протягом певного часу t дорівнює *0,75*. Заміна будь-якої пошкодженої лінії може бути проведена лише після пошкодження всіх ліній. Скільки потрібно провести ліній, щоб ймовірність функціонування зв'язку між пунктами протягом часу t була більша ніж *0,95*?

13.Відомо, що ймовірність переходу студента КІБС з першого курсу на другий дорівнює *0,85*, а ймовірність закінчити інститут – *0,78*. З якою ймовірністю можна стверджувати, що студент другого курсу закінчить інститут?

14. Ймовірність того, що банкомат пропрацює 10 000 годин, дорівнює 0,79, ймовірність того, що він пропрацює в два рази більше0,70. Банкомат пропрацювала 10 000 годин. Яка ймовірність того, що він пропрацює ще 10 000 годин?

15.В банківському відділенні працюють сім чоловіків і чотири жінки. За табельними номерами навмання відібрано три особи. Знайти ймовірність того, що всі відібрані особи є чоловіками.

16.На студента чекають іспити з економіки та банківської справи. На його думку ймовірність того, що він складе іспит з економіки, дорівнює 0,85, ймовірність того, що він здасть принаймні один предмет, дорівнює 0,96. А ймовірність того, що здасть обидва предмети – лише 0,78. Яка ймовірність того, що він складе іспит з банківської справи?

17.Студент КІБС знає 35 з 48 питань програми з економіки. Знайти ймовірність того, що студент знатиме запропоновані йому екзаменатором три питання з програми курсу.

18. Шість факультетів ВУЗу беруть участь у конкурсі на проведення науково-дослідницьких робіт. Ймовірність отримати контракт для кожного з них складають відповідно 0,3; 0,2; 0,5 і 0,1. Стало відомо, що перший факультет не приймає участі у конкурсі. Як змінилися шанси інших трьох факультетів?

19. В складі журі з економіки – три особи. Два його члени незалежно один від одного приймають вірне рішення з ймовірністю p, а третій для винесення рішення підкидає монету (кінцеве рішення виноситься однаковою кількістю голосів). Журі з однієї людини приймає вірне рішення з ймовірністю p. Яке з цих журі приймає вірне рішення з більшою ймовірністю?

20.Відомо, що в деякому регіоні 83 % фірм мають у штаті юриста і 52% компаній мають у штаті аналітиків. Вважається, що ці події незалежні. Яка ймовірність того, що фірма має одночасно в штаті і юриста і аналітика.

21.Серед 34 фірм, з яких 15 державних, а решта – приватні, приймають участь у тендері на 6 урядових контрактів. Вважається, що кожна фірма має рівні шанси на отримання контракту і за умовою, не може одержати більше одного контракту. Знайдіть імовірність, що принаймні одна державна фірма отримає контракт.

22.Ймовірність комівояжеру реалізувати товар, дорівнює 0,37. Яка ймовірність, що три спроби продажі в день приведуть до здійснення принаймні однієї покупки?

23.Фірма має можливість отримати два контракти. Ймовірність отримання першого контракту складає 0,9, а другого – 0,8. Вважаючи ці події незалежними, визначіть імовірність таких подій. Фірма отримає: а) A – обидва контракти, б) B – хоча б один контракт, в) C – жодного контракту, г) D – тільки один контракт, д) E – не більш одного контракту.

24. Ймовірність того, що в цеху буде в наявності сировина першого, другого, третього і четвертого типів відповідно дорівнюють 0,9; 0,8; 0,7; 0,6. Знайдіть імовірність того, що цех буде працювати, якщо для цього потрібно, щоб була в наявності сировина:

а) першого або другого і третього або четвертого типів;

б) і першого, і другого, і третього або четвертого типів;

в) першого і другого або третього або четвертого типів.

25.Серед 25 фірм, з яких 10 українських, розігрують 5 урядових контрактів. Вважається, що кожна фірма має рівні шанси на отримання контракту. За умовою фірма не може одержати більше одного контракту. Знайдіть імовірність, що принаймні одна українська фірма виграє контракт.

Глава 3. Формула повної ймовірності. Формула Байєса

§1. *Формула повної ймовірності*

Припустимо, що подія *A* може відбутися при умові появи однієї і тільки однієї з несумісних подій $H_1, H_2, \ldots, H_n$. Події цієї групи називають ***гіпотезами.***

Нехай відомі ймовірності гіпотез $H_1, H_2, \ldots, H_n$ і умовні ймовірності $P_1(A/H_1),\ P_2(A/H_2), \ldots P_n(A/H_n)$. Як визначити ймовірність події A в даній ситуації? Відповідь на це питання дає теорема.

Теорема 10. (*Формула повної ймовірності*). ***Ймовірність події А, яка може відбутися лише при умові появи однієї з несумісних подій – гіпотез*** $H_1, H_2, \ldots, H_n$***, що утворюють повну групу дорівнює сумі добутків ймовірностей гіпотез на відповідні умовні ймовірності події А:***

$$P(A) = P(H_1)P_1(A/H_1) + P(H_2)P_2(A/H_2) + \ldots + P(H_n)P_n(A/H_n) = \sum_{i=1}^{n} P(H_i)P_i(A/H_i). \qquad (11)$$

Доведення. За умовою подія *A* може відбутися, якщо відбудеться одна з несумісних подій $H_1, H_2, \ldots, H_n$, тобто поява події А означає появу однієї з несумісних подій $H_1 \cap A, H_2 \cap A, \ldots, H_n \cap A$ (незалежно якої). Це означає, що подія *A* є сумою несумісних подій $H_1 \cap A,\ H_2 \cap A,\ \ldots,\ H_n \cap A$:

$$A = H_1 \cap A \cup H_2 \cap A \cup \ldots \cup H_n \cap A.$$

Користуючись теоремою додавання ймовірностей несумісних подій запишемо:

$$P(A) = P(H_1 \cap A) + P(H_2 \cap A) + \ldots + P(H_n \cap A)$$

За теоремою множення ймовірностей залежних подій обчислюємо кожний доданок попередньої суми:

$$P(H_1 \cap A) = P(H_1) \cdot P_1(A/H_1);$$
$$P(H_2 \cap A) = P(H_2) \cdot P\ (A/H_2);$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$P(H_n \cap A) = P(H_n) \cdot P_n(A/H_n).$$

Додавши праві частини цих рівностей, одержимо:

$$P(A) = P(H_1)P_1(A/H_1) + P(H_2)P_2(A/H_2) + \ldots + P(H_n)P_n(A/H_n)$$

Теорема доведена.

Формула повної ймовірності застосовується в тих випадках, коли деяка подія *A* може відбутися в результаті появи однієї або кількох гіпотез $H_1, H_2, \ldots, H_n$ і потрібно обчислити її „повну” ймовірність із врахуванням всіх гіпотез.

Приклад 29. Банк кредитує три галузі промисловості: автомобільну. металургійну, вуглевидобувну в обсягах обсягах: 50%, 30%, 20% відповідно. Ймовірність неповернення кредитів рівна 4%, 8%, 6%. Який середній відсоток повернення кредитів по банку.

Розв’язування. Розглянемо гіпотези**:**

$H_1 = \{$ кредитується автомобільна галузь$\}$;

$H_2 = \{$ кредитується металургійна галузь$\}$;

$H_3 = \{$ кредитується вуглевидобувна галузь$\}$.

Очевидно, система подій - гіпотез H_1, H_2, H_3 утворює повну групу. Згідно умови задачі їх ймовірності дорівнюють. $P(H_1) = 0{,}5;\quad P(H_2) = 0{,}30;\quad P(H_3) = 0{,}20$.

Для відшуканні ймовірності події $A = \{$навманя взятий кредит повернеться$\}$ використаємо формулу повної ймовірності (11):

$$P(A) = P(H_1)P\ (A/H_1) + P(H_2)P\ (A/H_2) + P(H_3)P\ (A/H_3) =$$
$$= 0{,}5 \cdot 0{,}96 + 0{,}30 \cdot 0{,}92 + 0{,}20 \cdot 0{,}94 = 0{,}944.$$

Приклад 30. В ящику знаходяться 1000 деталей: 700 з однієї партії, 300 з другої. Визначіть ймовірність того, що навмання вибрана деталь з ящику виявиться бракованою, якщо ймовірність браку в першій партії дорівнює 0,06, а в другій 0,04.

Розв'язування. Нехай подія A – навмання вибрана деталь виявилась бракованою. Можливі наступні припущення (гіпотези) про те, до якої партії належить дана деталь: H_1 – деталь з першої партії, H_2 – деталь з другої партії. Ймовірності цих гіпотез дорівнюють: $P(H_1) = \frac{700}{1000} = 0{,}7$, $P(H_2) = \frac{300}{1000} = 0{,}3$. Умовні ймовірності відомі з умови задачі $P(A/H_1) = 0{,}06$, $P(A/H_2) = 0{,}04$. Шукану ймовірність знаходимо за формулою повної ймовірності

$$P(A) = P(H_1)\cdot P(A/H_1) + P(H_2)\cdot P(A/H_2) = 0{,}7\cdot 0{,}06 + 0{,}3\cdot 0{,}04 = 0{,}054.$$

Приклад 31. В двох урнах знаходяться білі та чорні кулі. Склад першої урни – 14 білих та 16 чорних, другої – 25 білих та 10 чорних куль. З кожної урни виймають навмання по одній кулі, а потім із цих двох куль навмання обирають одну кулю. Знайти ймовірність того, що ця куля – біла?

Розв'язування. Введемо подію $A = \{$вибрана біла куля$\}$. Зазначимо, що результат вибору однієї кулі із двох залежить від того, якого кольору будуть дві кулі. Введемо гіпотези:

$H_1 = \{$з першої урни вийнято *білу* кулю, з другої урни - *білу* кулю$\}$;

$H_2 = \{$з першої урни вийнято *білу* кулю, з другої урни - *чорну* кулю$\}$;

$H_3 = \{$з першої урни вийнято *чорну* кулю, з другої урни - *білу* кулю$\}$;

$H_4 = \{$з першої урни вийнято *чорну* кулю, з другої урни - *чорну* кулю$\}$.

Оскільки кулі виймаються із урн незалежно одна від одної, то ймовірності гіпотез відповідно дорівнюють:

$$P(H_1) = \frac{14}{30}\cdot\frac{25}{35} = \frac{1}{3};\qquad P(H_2) = \frac{14}{30}\cdot\frac{10}{35} = \frac{2}{15};$$
$$P(H_3) = \frac{16}{30}\cdot\frac{25}{35} = \frac{8}{21};\qquad P(H_4) = \frac{16}{30}\cdot\frac{10}{35} = \frac{16}{105}.$$

Знайдемо умовні ймовірності:

$$P_1(A/H_1)=0;\ P(A/H_2)=P(A/H_3)=\frac{1}{2};\quad P(A/H_4)=1.$$

Далі, згідно формули повної ймовірності (11), одержимо

$$P(A)=P(H_1)P(A/H_1)+P(H_2)P(A/H_2)+P(H_3)P(A/H_3)+P(H_4)P(A/H_4)=$$
$$=\frac{1}{3}\cdot 0+\frac{2}{15}\cdot\frac{1}{2}+\frac{8}{21}\cdot\frac{1}{2}+\frac{16}{105}\cdot 1=\frac{43}{105}.$$

§2. *Ймовірність гіпотез. Формули Байєса*

Розглянемо повну групу несумісних подій-гіпотез $H_1, H_2, \ldots, H_n$, ймовірності яких $P(H_i)$, $(i=1,2,..n)$ відомі до проведення випробування. Подія А може відбуватися при умові появи однієї з описаних подій - гіпотез. Проводиться випробування, в результаті якого зареєстрована поява події A, причому відомо, що цій події наші гіпотези приписували певні ймовірності $P(A/H_i),(i=1,2,...n)$. Виникає питання, які будуть ймовірності цих гіпотез після проведення випробування.

Відкидаємо гіпотези, які заперечують подію A і обчислюємо умовні ймовірності $P(A/H_i),(i=1,2,...n)$.

На основі теореми множення ймовірностей залежних подій маємо.

$$P(A\cdot H_i)=P(A)\cdot P(H/A_i)=P(H_i)\cdot P(A/H).$$

Звідси знайдемо

$$P(H/A_i)=\frac{P(H_i)\cdot P(A/H_i)}{P(A)}.$$

З формули (11) повної ймовірності

$$P(A)=\sum_{j=1}^{n}P(H_j)\cdot P_j(A/H_i)$$

для ймовірності $P(H/A_i)$ одержимо

$$P(H/A_i)=\frac{P(H_i)\cdot P(A/H_i)}{\sum\limits_{j=1}^{n}P(H_j)\cdot P(A/H_i)},\ (i=1,2,..n). \tag{12}$$

Останнє співвідношення називається ***формулою Байєса***.

Приклад 32. В умовах попередньої задачі 30 знайдіть ймовірність того, що навмання вибрана деталь належить першій (другій) партії, якщо вона виявилась бракованою.

Розв'язування. Шукані ймовірності за формулою Байєса дорівнюють:

$$P(H_1/A)=\frac{P(H_1)\cdot P(A/H_1)}{P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)}=\frac{0,7\cdot 0,06}{0,054}=\frac{7}{9},$$

$$P(H_2/A)=\frac{P(H_2)\cdot P(A/H_2)}{P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)}=\frac{0,3\cdot 0,04}{0,054}=\frac{2}{9}.$$

Приклад 33. Позики, які надає банк, можуть бути перевірені на предмет їх ризикованості одним з двох кредитних інспекторів. Ймовірність того, що позика буде перевірена першим інспектором, дорівнює *0,7*, другим – *0,3*. Ймовірність того, що позика буде визнана не ризикованою першим інспектором, дорівнює 0,95, другим – 0,98. Позика в результаті перевірки визнана не ризикованою. Знайти ймовірність того, що позику перевірив перший інспектор.

Розв'язування. До випробування можливі дві гіпотези: H_1 і H_2 - позикаь буде перевірена відповідно першим або другим інспектором, причому $P(H_1)=0{,}7$, $P(H_2)=0{,}3$.

Умовні ймовірності події $A=\{$позика не ризикована$\}$ дорівнюють:

$$P\left(A/H_1\right)=0{,}95;\; P\left(A/H_2\right)=0{,}98$$

Для обчислення ймовірності $P\left(H_1/A\right)$ використаємо формулу (12) і отримаємо:

$$P\left(H_1/A\right)=\frac{P(H_1)\cdot P\left(A/H_1\right)}{P(H_1)\cdot P\left(A/H_1\right)+P(H_2)\cdot P\left(A/H_2\right)}=\frac{0{,}7\cdot 0{,}95}{0{,}7\cdot 0{,}95+0{,}3\cdot 0{,}98}\approx 0{,}693$$

Отже, ймовірність того, що не ризикована позика перевірена першим контролером, дорівнює 0,693.

Завдання для самостійної роботи студентів

1. Є два набори дискет. Імовірність того, що дискета з першого набору працює, дорівнює 0,8, з другого – 0,9. Знайдіть імовірність того, що навмання взята дискета (з навмання взятого набору) працює.

2. Інвестор вкладає свої гроші в акції великого ризику, в акції сприйнятливого ризику та безризикові акції в пропорціях 10 %, 30 % і 60 %. Імовірність отримання прибутку від цих акцій складає 0,6; 0,75 і 0,9 відповідно. Знайдіть: а) ймовірність того, що інвестор отримав прибуток; б) якщо інвестор отримав прибуток, то, яка ймовірність, що він отриманий від безризикових акцій?

3. Лиття в болванках поступає з двох заготівельних цехів: 70 % – з першого, 30 % – з другого. Причому матеріал з першого цеху має 10 % браку, з другого – 20 %. Знайдіть імовірність того, що: а) навмання взята болванка з дефектом, б) болванка виготовлена першим цехом, якщо вона виявилася з дефектом.

4. Фірма збирається випускати новий товар на ринок. Підраховано, що ймовірність доброго збуту продукції на ринок дорівнює 0,6; поганого – 0,4. Компанія збирається провести маркетингове дослідження, ймовірність правильності якого 0,8. Як змінюються початкові ймовірності рівня реалізації, якщо це дослідження передбачає поганий збут?

5.У продаж надійшли телевізори з трьох заводів. Продукція першого заводу становить 30 %, другого – 20 %, третього – 50 %. Продукція першого заводу містіть у собі 20 % телевізорів з прихованим дефектом, другого – 10 %, третього – 5 %. Яка ймовірність того, що: а) навмання куплений телевізор буде якісним, б) телевізор виготовлений на першому заводі, якщо він виявився якісним?

6. Припустимо, що 5 % чоловіків та 0,25 % жінок хворіють на дальтонізм. Навмання обрана людина хворіє на дальтонізм. Яка ймовірність того, що це чоловік?

7. В урні дві білих і чотири чорних кулі. Навмання взяли 2 кулі. Після цього – третю. Знайдіть імовірність того, що третя куля – біла.

8.На екзамен запропоновано 30 білетів, з них 25 простих і 5 складних. Яка ймовірність того, що студент, зайшовши на іспит другим, витягне простий білет?

9.П'ятнадцять екзаменаційних білетів містять по два питання. Студент вивчив лише 25 питань. Знайдіть імовірність того, що студент складе іспит, якщо для цього досить відповісти на обидва питання одного білета або на одне питання з першого білета і на додаткове питання з другого білета.

13. На трьох виробничих лініях виготовляються однакові вироби. Відомо, що 25% всієї продукції виробляється першою лінією, 35% - другою, і 40% - третьою. Ймовірність виготовлення виробу, що відповідає стандарту, першою лінією дорівнює 0,97, другою – 0,96 і третьою – 0,98. Виготовлена протягом дня продукція знаходиться на складі. Знайти ймовірність того, що навмання взятий виріб не стандартний.

14. Відносні частоти помилок у роботі комп'ютера відносяться як 2:4:7. Ймовірність виявлення неполадок у роботі мікропроцесора, оперативної пам'яті та решти пристроїв комп'ютера відповідно дорівнюють 0,8; 0,7; і 0,9. Знайти ймовірність неполадок у роботі комп'ютера.

15. На складання автомобіля поступають деталі з трьох автоматів. Відомо, що перший автомат робить 0,3 % браку, другий – 0,2 %, третій – 0,4 %. Знайдіть імовірність того, що на складання надійшла бракована деталь, якщо з першого автомату надійшло – 1000 деталей, з другого – 2000, з третього – 2500.

16. У першій коробці 20 ламп, з них 18 працюють, у другій – 10 ламп, з них 9 працюють. З другої коробки взяли лампу і поклали в першу. Знайдіть імовірність того, що навмання взята лампа з першої коробки працює.

17. В цеху працюють 20 верстатів. З них 10 марки А, 6 марки В, 4 марки С. Імовірність того, що верстати А, В, С виготовляють деталь відмінної якості, відповідно дорівнює 0,9; 0,8; та 0,7. Який відсоток деталей відмінної якості виготовляють у цьому цеху?

18. На роботу менеджера претендують 45% жінок і 55% чоловіків. Серед жінок 38% мають університетську освіту, а серед чоловіків-60%. Знайти:

а) ймовірність того, що вибрана навмання заява є від особи з університетською освітою?

б) ймовірність того, що ця особа з університетською освітою є жінка?

в) ймовірність того, що вибрана навмання заява є від жінки з університетською освітою?

19. В групі 28 студентів, з яких 8 знають 93% екзаменаційних питань по кожному з трьох розділів програми, 11- 67%, 9-76% і 8-55%. Знайти ймовірність того, що:

а) на іспиті студент з цієї групи дав вірні відповіді на два питання з перших двох розділів програми і відмовився відповідати на питання третього розділу;

б) студент, який відповідав згідно пункту а), вивчив 95% програми?

20. Завод виготовляє вироби, кожен з яких з ймовірність 0,008 має дефект. В цеху виріб з рівною ймовірністю перевіряється одним з двох контролерів. Перший контролер визначає наявний дефект з ймовірністю 0,72, а другий – з ймовірністю 0,85.

Якщо в цеху виріб не забраковано, то він надходить у відділ технічного контролю заводу, де дефект, якщо він є, визначається з ймовірністю 0,87. Відомо, що виріб забраковано. Знайти ймовірність того, що він забракований: а) першим контролером;б) відділом технічного контролю заводу.

21.Два бухгалтери заповнюють документи, які складають у спільну папку. Ймовірність появи помилки в документі першого бухгалтера дорівнює 0,19, а другого – 0,2. Перший бухгалтер заповнив 45 документів, а другий – 55. У навмання вибраному з папки документі виявилася помилка. Визначити ймовірність того, що його склав перший бухгалтер.

22.В аудиторській фірмі працює три аудитора і частки розглядуваних ними справ відносяться як 6:8:7. Ймовірність проведення вірного аудиту у кожного з аудиторів відповідно дорівнює 0,84, 0,76 та 0,63. Клієнт фірми отримав вірний аудит. Знайти ймовірність того, що його аудитором був перший аудитор.

23.Ймовірності попадання в ціль при одному пострілі кожним із стрільців відповідно дорівнюють 0,7;0,45 та 0,82. Один з трьох стрільців викликається на лінію вогню, робить два постріли і в ціль не попадає. Знайти ймовірність того, що постріли виконував перший стрілець.

Глава 4. Повторні незалежні випробування. Схема Бернуллі.

§1. Послідовність незалежних випробувань, формула Бернуллі.

Припустимо, що проводиться послідовно декілька випробувань (експериментів), кожний з яких має лише два несумісні наслідки (випадкові події A і $\overline{A}$).

Означення. ***Випробування називаються незалежними відносно події А , якщо ймовірність цієї події в кожному випробуванні не залежить від результатів інших випробувань.***

Означення. ***Серія повторних незалежних випробувань, в кожному з яких випадкова подія А має одну і ту ж ймовірність $P(A)=p$, яка не залежить від випробування, називається схемою Бернуллі.***

За схемою Бернуллі в кожному випробуванні можливі тільки два наслідки: 1) відбудеться подія A і 2) відбудеться подія $\overline{A}$ зі сталими ймовірностями $P(A)=p$ і $P(\overline{A})=1-p=q$ відповідно, причому $p+q=1$.

Розглянемо задачу: у випадку схеми Бернуллі визначити ймовірність $P_n(m)$ $(0 \le m \le n)$ того, що при n випробуваннях випадкова подія A, яка має одну і ту ж ймовірність $P(A)=p$ в кожному окремому випробуванні, з'явиться рівно m разів.

Зауважимо, що при цьому не вимагається, щоб подія повторилася рівно m разів в певній послідовності. Так, якщо мова йде про появу події A тричі в трьох випробуваннях, то можливі такі складні події: $A \cap A \cap \overline{A},\ A \cap \overline{A} \cap A,\ \overline{A} \cap A \cap A,\ A \cap A \cap A$. Запис $A \cap A \cap \overline{A}$ означає, що в першому і другому випробуваннях подія A з'явилася, а в третьому випробуванні вона не з'явилася, тобто з'явилася протилежна подія $\overline{A}$.

Справедливе наступне **твердження**: ***ймовірність того, що в результаті n випробувань (експериментів) за схемою Бернуллі випадкова подія А з'явиться m разів,*** *обчислюється за формулою*

$$P_n(m) = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m} \qquad (13)$$

Дійсно, ймовірність складної події, яка полягає в тому, що в випробуваннях випадкова подія A з'явиться m разів і не з'явиться $n-m$ разів за теоремою множення ймовірностей незалежних подій дорівнює:

$$\underbrace{P(A)\cdot \ldots \cdot P(A)}_{\text{m разів}} \cdot \underbrace{P(\overline{A})\cdot \ldots \cdot P(\overline{A})}_{\text{n-m разів}} = p^m q^{n-m}$$

Число таких складних подій співпадає з числом комбінацій з n елементів по m елементів, тобто дорівнює C_n^m. Оскільки ці складні події несумісні, то за теоремою додавання ймовірностей несумісних подій шукана ймовірність дорівнює сумі ймовірностей всіх можливих складних подій. Враховуючи, що ймовірності всіх цих складних подій є однакові $\left(p^m q^{n-m}\right)$, то шукана ймовірність дорівнює добутку ймовірності однієї складної події на їх число C_n^m, тобто

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

Формула (13) називається ***формулою Бернуллі (біноміальною).***

Окремі випадки застосування формули Бернуллі

1. Ймовірність того, що подія А з'явиться в усіх випробуваннях дорівнює:

$$P_n(n) = C_n^n p^n q^{n-n} = p^n. \qquad (14)$$

2. Ймовірність того, що в n випробуваннях подія А не з'явиться жодного разу дорівнює:

$$P_n(0) = C_n^0 p^0 q^{n-0} = q^n. \qquad (15)$$

3. Ймовірність того, що в n випробуваннях подія А з'явиться не менше m разів, дорівнює:

$$P\ (A \ge m) = P_n(m) + P_n(m+1) + \ldots + P_n(n). \qquad (16)$$

4. Ймовірність того, що в n випробуваннях подія А з'явиться не менше m_1 разів і не більше m_2 разів, дорівнює:

$$P\ (m_1 \le A \le m_2) = P_n(m_1) + P_n(m_1+1) + \ldots + P_n(m_2). \quad (17)$$

Зауваження. Для всіх **n** випробувань по схемі Бернуллі виконується рівність:

$$P_n(0) + P_n(1) + \ldots + P_n(m) + \ldots + P_n(n) = 1 \quad .$$

Приклад 34. Із статистики відомо, що 90% сімей мають принаймні один автомобіль. Яка ймовірність того, що три з чотирьох навмання вибраних сімей мають автомобіль?

Розв'язування. Нехай подія А={одна сім'я має автомобіль}. Очевидно $P(A) = p = 0{,}9, q = 0{,}1$. Шукана ймовірність є ймовірністю того, що випадкова подія А з'явиться тричі у чотирьох незалежних повторних випробуваннях. За формулою Бернуллі (12)

$$P_4(3) = C_4^3 \cdot 0{,}9^3 \cdot 0{,}1 = \frac{4!}{3!\cdot 1!} \cdot 0{,}9^3 \cdot 0{,}1 = 4 \cdot 0{,}9^3 \cdot 0{,}1 \approx 0{,}3.$$

Приклад 35. Банк запровадив 9 нових кредитних продуктів (позик). Ймовірність того, що запровадження одного виду позики дасть прибуток, дорівнює 0,8. Запровадження вважається успішним, якщо протягом дня буде заключено не менше семи угод. Яка ймовірність успіху, якщо угоди заключаються незалежно.

Розв'язування. Нехай подія А={заключення однієї угоди дає прибуток}. За умовою задачі $p = P(A) = 0{,}8, q = P(\overline{A}) = 0{,}2$. Успіх визначається тим, що буде заключено не менше семи угод. Ймовірність успіху $P_9(7 \le m \le 9)$ є сума ймовірностей $P_9(7), P_9(8), P_9(9)$, тобто

$$P_9(7 \le m \le 9) = P_9(7) + P_9(8) + P_9(9).$$

Обчислимо ймовірності:

$$P_9(7)=C_9^7\cdot 0{,}8^7\cdot 0{,}2^2=\frac{9!}{7!\cdot 2!}0{,}8^7\cdot 0{,}2^2=4\cdot 9\cdot 0{,}8^7\cdot 0{,}2^2\approx 0{,}302;$$

$$P_9(8)=C_9^8\cdot 0{,}8^8\cdot 0{,}2=\frac{9!}{8!\cdot 1!}0{,}8^8\cdot 0{,}2=9\cdot 0{,}8^8\cdot 0{,}2\approx 0{,}302;$$

$$P_9(9)=C_9^9\cdot 0{,}8^9\cdot 0{,}2^0=0{,}8^9\approx 0{,}134.$$

Отже, ймовірність успіху кредитних проектів на ринку позик дорівнює

$$P_9(7\le m\le 9)\approx 0{,}302+0{,}302+0{,}134=0{,}738.$$

§2. *Найімовірніше число появи випадкової події*.

Розглянемо приклад:

Приклад 36. Ймовірність появи випадкової події A в кожному з $n=6$ незалежних випробувань за схемою Бернуллі є величиною сталою і дорівнює $p=0{,}5$. Обчислимо і проаналізуємо ймовірності складних подій для $m=0,1,2,3,4,5,6$

Розв'язування. Обчислимо ймовірності за формулою Бернуллі (13)

$$P_6(0)=C_6^0\left(\frac{1}{2}\right)^0\cdot\left(\frac{1}{2}\right)^6=\frac{1}{256},\qquad P_6(4)=C_6^4\left(\frac{1}{2}\right)^4\cdot\left(\frac{1}{2}\right)^2=\frac{15}{256},$$

$$P_6(1)=C_6^1\left(\frac{1}{2}\right)^1\cdot\left(\frac{1}{2}\right)^5=\frac{6}{256},\qquad P_6(5)=C_6^5\left(\frac{1}{2}\right)^5\cdot\left(\frac{1}{2}\right)^1=\frac{6}{256},$$

$$P_6(2)=C_6^2\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)^4=\frac{15}{256},\qquad P_6(6)=C_6^6\left(\frac{1}{2}\right)^6\cdot\left(\frac{1}{2}\right)^0=\frac{1}{256},$$

$$P_6(3)=C_6^3\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^3=\frac{20}{256}.$$

За результатами обчислень складемо таблицю:

m	0	1	2	3	4	5	6
$P_n(m)$	$\frac{1}{256}$	$\frac{6}{256}$	$\frac{15}{256}$	$\frac{20}{256}$	$\frac{15}{256}$	$\frac{6}{256}$	$\frac{1}{256}$

З наведеної таблиці бачимо, що ймовірність $P_n(m)$ набуває найбільшого значення при $m=3$, а саме $P_6(3)=\frac{20}{256}$. Отже, найімовірніше число появи події A є $m_0=3$.

Тепер розглянемо загальний випадок.

Означення. ***Найімовірнішим числом появи випадкової події*** A ***в незалежних випробуваннях за схемою Бернуллі називається число*** m_0***, для якого ймовірність перевищує, або в усякому разі не є меншою, за ймовірність решти можливих наслідків випробувань (експериментів).***

У наведеному вище прикладі для визначення найімовірнішого числа m_0 появи події A ми обчислювали ймовірності для різних можливих значень m ($0 \le m \le n$). Визначимо число m_0.

Для цього, обчислимо ймовірності $P_n(m)$ для $m = m_0$, $m = m_0 - 1$, $m = m_0 + 1$ і розглянемо їх відношення:

$$\frac{P_n(m_0)}{P_n(m_0-1)} \ge 1 \Leftrightarrow \frac{C_n^{m_0} p^{m_0} q^{n-m_0}}{C_n^{m_0-1} p^{m_0-1} q^{n-m_0+1}} \ge 1 \Leftrightarrow$$

$$\Leftrightarrow \frac{\dfrac{n!}{m_0!(n-m_0)!} p^{m_0} q^{n-m_0}}{\dfrac{n!}{(m_0-1)!(n-m_0+1)!} p^{m_0-1} q^{n-m_0+1}} \ge 1 \Leftrightarrow$$

$$\Leftrightarrow \frac{\dfrac{p}{m_0}}{\dfrac{q}{n-m_0+1}} \ge 1 \Leftrightarrow \frac{p(n-m_0+1)}{qm_0} \ge 1 \Leftrightarrow m_0 \le np + p$$

Аналогічно одержимо, що

$$\frac{P_n(m_0)}{P_n(m_0+1)} \ge 1 \Leftrightarrow m_0 \ge np - q$$

Об’єднавши кінцеві нерівності , дістанемо:

$$np - q \le m_0 \le np + p \qquad (18)$$

З одержаної нерівності (18) визначається натуральне число m_0.

Приклад 37. Ймовірність несвоєчасної сплати кредиту для кожного з 500 підприємців міста (подія А) дорівнює $p = 0{,}1$. Знайти найімовірнішу кількість підприємців міста, котрі вчасно не повернуть позику, і визначити відповідну ймовірність.

Розв’язування. За умовою задачі $n = 500$, $p = 0{,}1$, $q = 0{,}9$. Найімовірніше число підприємців m_0, котрі не вчасно повернуть позику, задовольняє нерівність

$$500 \cdot 0{,}1 - 0{,}9 \le m_0 \le 500 \cdot 0{,}1 + 0{,}1 \Leftrightarrow 49{,}1 \le m_0 \le 50{,}1 \Rightarrow m_0 = 50$$

§3. *Локальна теорема Лапласа*

Якщо число випробувань n велике, то обчислення ймовірності $P_n(m)$ за формулою Бернуллі приводить до громіздких обчислень. Лаплас довів важливу наближену формулу для обчислення ймовірності $P_n(m)$ появи випадкової події A рівно m разів при досить великому числі n випробувань.

Теорема (*локальна теорема Муавра-Лапласа*). *Нехай* $p = P(A)$***- ймовірність випадкової події*** A***, причому*** $0 < p < 1$***. Тоді ймовірність того, що в схемі Бернуллі подія*** A ***з’явиться рівно*** m ***разів при*** n ***випробуваннях, виражається наближеною формулою:***

$$P_n(m) \approx \frac{\varphi(x)}{\sqrt{npq}}\, при\, x = \frac{m-np}{\sqrt{npq}}, \tag{19}$$

де $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ — ***функцією Гаусса.***

Встановлено, що відносна похибка формули (19) прямує до нуля, при $n \to \infty$. З доведенням даної теореми можна ознайомитися в рекомендованій літературі [1].

Функція Гаусса табульована (див. Додаток 1) для додатних значень x. Для від'ємних значень аргумента x користуються тією ж таблицею, оскільки функція $\varphi(x)$ парна.

Функція $\varphi(x)$ має широкі застосування в теорії ймовірностей та математичній статистиці. Наведемо основні властивості функції Гаусса:

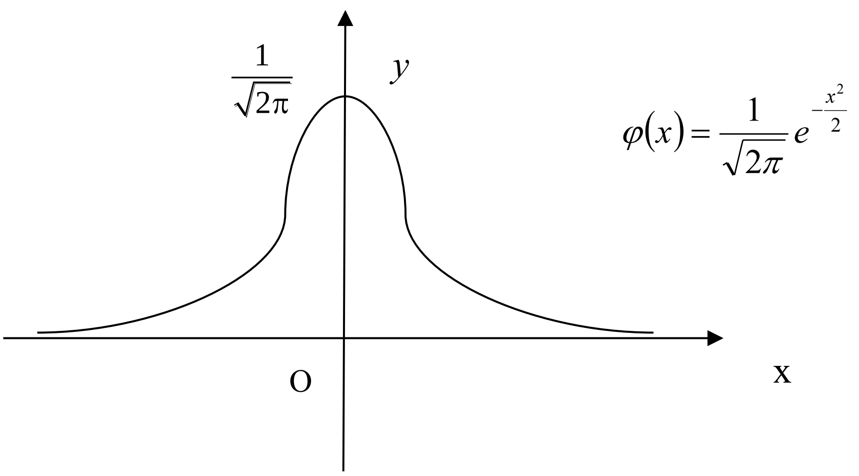
1) $\varphi(x)$ визначена для всіх $x \in (-\infty, \infty)$ і $\varphi(x) > 0$;
2) $\varphi(x)$ є парною функцією, тобто $\varphi(-x) = \varphi(x)$;
3) $\lim\limits_{x \to \pm\infty} \varphi(x) = 0$;
4) $\varphi(x)$ має максимум в точці $x = 0$, тобто $\max \varphi(x) = \varphi(0) = \frac{1}{\sqrt{2\pi}}$.

Дійсно:

$\varphi'(x) = -x\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Rightarrow \varphi'(0) = 0 \Rightarrow x = 0$ - критична точка екстремума функції $\varphi(x)$. Далі, $\varphi'(x)\big|_{x<0} > 0$ і $\varphi'(x)\big|_{x>0} < 0 \Rightarrow$ у критичній точці $x = 0$ функція $\varphi(x)$ має максимум;

5) $\varphi''(x) = (x^2 - 1)\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Rightarrow \varphi''(\pm 1) = 0 \Rightarrow x = \pm 1$ - критичні точки перегину графіка функції $\varphi(x)$. При цьому, $\varphi''(x)\big|_{x<-1} > 0, \varphi''(x)\big|_{-1<x<1} < 0, \varphi''(x)\big|_{x>1} > 0$. Отже, графік функції $\varphi(x)$ вгнутий при $x \in (-\infty, -1) \cup (1, \infty)$ і опуклий при $x \in (-1, 1)$.

Графік функції Гаусса зображений на мал 3.



Мал. 3.

Зауважимо, що функція $\varphi(x)$ протабульована для значень аргумента $0 \le x \le 3{,}9$, а для значень $x \ge 4$ приймають, що $\varphi(x) = 0$ (аналогічно, $\varphi(x) = 0$ при $x \le -4$).

Приклад 38. В банку 25% вкладників мають вклади, суми яких перевищують 100000 грн. Навмання обирають 100 вкладників. Знайти ймовірність того, що в їх числі виявиться:

а) 70, б) 30, в) 20 вкладників, вклади яких перевищують 100000 грн.

Розв'язування. За умовою задачі маємо, що $n = 100,\ p = 0{,}25,\ q = 0{,}75,\ m = 70{,}30{,}20$

Шукані ймовірності $P_{100}(70), P_{100}(30), P_{100}(20)$ обчислюємо за формулою Лапласа (19), бо застосування формули Бернуллі (13) приводить до певних громіздких обчислень.

$$\sqrt{npq} = \sqrt{100 \cdot 0{,}75 \cdot 0{,}25} = \sqrt{18.75} \approx 4{,}33$$

а)
$$x = \frac{m - np}{\sqrt{npq}} = \frac{70 - 100 \cdot 0{,}25}{4{,}33} = \frac{45}{4{,}33} \approx 10{,}39;$$

$$P_{100}(70) = \frac{\varphi(10{,}39)}{4{,}33} \approx 0{,}00001.$$

б)
$$x = \frac{30 - 100 \cdot 0{,}25}{4{,}33} = \frac{5}{4{,}33} \approx 1{,}155;$$

$$P_{100}(30) = \frac{\varphi(1{,}155)}{\sqrt{npq}} = \frac{0{,}2047}{4{,}33} \approx 0{,}0473.$$

в)
$$x = \frac{20 - 100 \cdot 0{,}25}{4{,}33} = \frac{-5}{4{,}33} \approx -1{,}1547;$$

$$P_{100}(20) = \frac{\varphi(-1{,}1547)}{4{,}33} = \frac{0{,}2083}{4{,}33} \approx 0{,}048.$$

§4. *Інтегральна теорема Лапласа*

Знайдемо тепер ймовірність $P_n(m_1, m_2)$ того, що в схемі Бернуллі випадкова подія, яка має ймовірність успіху $P(A) = p \quad (0 < p < 1)$ в кожному випробуванні, з'явиться від m_1 до m_2 разів. Відповіддю тут є *інтегральна теорема Лапласа*, яку сформулюємо нижче.

Теорема. ***Якщо ймовірність*** $P(A) = p \quad (0 < p < 1)$ ***в кожному випробуванні є однаковою, то ймовірність того, що в схемі Бернуллі подія*** A ***з'явиться від*** m_1 ***до*** m_2 ***разів при*** n ***випробуваннях, виражається наближеною формулою***

$$P_n(m_1, m_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{z^2}{2}} dz, \tag{20}$$

де $x_1 = \frac{m_1 - np}{\sqrt{npq}}, x_2 = \frac{m_2 - np}{\sqrt{npq}}$.

Формулу (20) записують за допомогою ***функції Лапласа (інтегралу ймовірностей)***:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz. \tag{21}$$

Тоді

$$P_n(m_1, m_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{0} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_0^{x_2} e^{-\frac{z^2}{2}} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{x_2} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_0^{x_1} e^{-\frac{z^2}{2}} dz = \Phi(x_2) - \Phi(x_1)$$

і формула (20) набуває вигляду:

$$P_n(m_1, m_2) \approx \Phi(x_2) - \Phi(x_1). \tag{22}$$

Формули (20) і (22) виражають наближено ймовірність того, що число m появи події A в n випробуваннях за схемою Бернуллі задовольняє нерівність $m_1 \le m \le m_2$.

Формулу (22) часто записують так:

$$P_n(m_1 \le m \le m_2) \approx \Phi(x_2) - \Phi(x_1). \tag{23}$$

її називають ***інтегральною формулою Лапласа.***

Функція Лапласа (21) також часто використовується у теорії ймовірностей і математичній статистиці і має такі основні властивості:

1) $\Phi(x)$ визначена для всіх $x \in (-\infty, \infty)$;

2) $\Phi(-x) = -\Phi(x)$, тобто $\Phi(x)$ є непарною функцією;

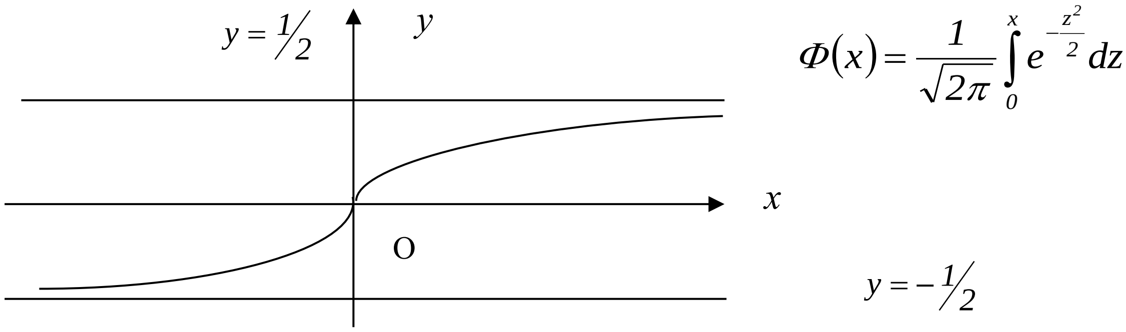
3) $\Phi(0) = 0; \Phi(\infty) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{z^2}{2}} dz = 0{,}5$, оскільки $\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}$ є інтеграл Пуассона; $\Phi(-\infty) = 0{,}5$, оскільки $\Phi(x)$ непарна функція;

4) $\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} > 0$, тобто $\Phi(x)$ є зростаючою функцією;

5) $\Phi''(x) = -x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Rightarrow \Phi''(x)\big|_{x<0} > 0$ і $\Phi''(x)\big|_{x>0} < 0 \Rightarrow$ графік $\Phi(x)$ вгнутий в інтервалі $(-\infty, 0)$ і опуклий в інтервалі $(0, \infty)$;

6) $\lim_{x \to -\infty} \Phi(x) = 0{,}5$ і $\lim_{x \to +\infty} \Phi(x) = 0{,}5 \Rightarrow$ прямі $y = -0{,}5$ і $y = 0{,}5$ є відповідно лівосторонньою і правосторонньою асимптотами графіка функції $\Phi(x)$.

Графік функції Лапласа зображений на мал. 4.



Мал. 4.

Інтеграл, який фігурує у формулі (21), не виражається через елементарні функції. Значення функції $\Phi(x)$ табульовані (див. Додаток 2) для значення аргумента x від 0 до 5. Для від'ємних значень x користуємося цією ж таблицею, бо $\Phi(-x) = -\Phi(x)$. Для значень аргумента x, ***для яких*** $|x| > 5$, ***дотримуються правила:*** $\Phi(x) = 0{,}5$.

Приклад 39. Ймовірність збанкрутіти під час економічної кризи підприємцю, який є клієнтом банку дорівнює 0,2. Яка ймовірність того, що серед 1000 підприємців клієнтів банку збанкрутує від 90 до 200 осіб?

Розв'язування. Нехай подія А={вибраний навмання підприємець збанкрутів}. За умовою задачі $p = P(A) = 0{,}2$, $q = 0{,}8$, $n = 1000$, $m_1 = 90$, $m_2 = 200$. Для обчислення шуканої ймовірності $P_n(m_1 \le m \le m_2)$ використаємо формулу (22).

Знайдемо:

$$\sqrt{npq} = \sqrt{1000 \cdot 0{,}2 \cdot 0{,}8} = \sqrt{160} \approx 12{,}65;$$

$$x_1 = \frac{90 - 1000 \cdot 0{,}2}{12{,}65} = \frac{-110}{12{,}65} \approx -8{,}6957; \quad x_2 = \frac{200 - 1000 \cdot 0{,}2}{12{,}65} = 0$$

За формулою (23) і таблицею Додатку 2 , одержимо:

$$P_{1000}(90 \le m \le 200) = \Phi(0) - \Phi(-8{,}6957) = 0 + 0{,}4999 = 0{,}4999.$$

§5.*Формула Пуассона*

При малих ймовірностях успіху p в окремому випробуванні застосовують інші наближені формули. Нехай проводиться серія з n незалежних випробувань (n=1,2,3,…), причому ймовірність появи випадкової події A в серії $P(A) = p_n$ залежить від номера n і $p_n \to 0$ при $n \to \infty$. Припустимо, що для кожної серії середнє значення числа появ події A є стале, тобто

$$n \cdot p_n = \mu = const.$$

Для кожного фіксованого n маємо, що

$$p_n = \frac{\mu}{n}.$$

На основі формули Бернуллі (13) для ймовірності $P_n(m)$ появи події рівно m разів в n випробуваннях маємо вираз:

$$P_n(m)=C_n^m p_n^m (1-p_n)^{n-m} = C_n^m\left(\frac{\mu}{n}\right)^m\left(1-\frac{\mu}{n}\right)^{m-n}$$

Зафіксуємо число m і перейдемо в рівності (13) до границі при $n\to\infty$. Для цього обчислимо границю множників у правій частині рівності.

$$\lim_{n\to\infty} C_n^m\left(\frac{\mu}{n}\right)^m = \lim_{n\to\infty}\frac{n!}{m!(n-m)!}\cdot\left(\frac{\mu}{n}\right)^m = \lim_{n\to\infty}\cdot\frac{n(n-1)(n-2)\ldots[n-(m-1)]}{m!n^m}\mu^m =$$

$$=\frac{\mu^m}{m!}\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{m-1}{n}\right)=\frac{\mu^m}{m!}$$

Для обчислення границі другого множника $\left(1-\frac{\mu}{n}\right)^{m-n}$ використаємо відому границю:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e \Rightarrow \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = e^{-1},$$

згідно якої

$$\left(1-\frac{\mu}{n}\right)^{n-m} = \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n\left(1-\frac{\mu}{n}\right)^{-m} = \lim_{n\to\infty}\left[\left(1-\frac{\mu}{n}\right)^{\frac{n}{\mu}}\right]^{\mu}\cdot\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^{-m} = \left(e^{-1}\right)^{\mu}\cdot 1 = e^{-\mu}.$$

В результаті з рівності (13) одержуємо, що

$$\lim_{n\to\infty} P_n(m) = \frac{\mu^m}{m!}e^{-\mu}.$$

Якщо n велике число, то різниця між ймовірністю $P_n(m)$ та її границею є як завгодно малою, тобто можемо записати наближену ***формулу Пуассона.***

$$P_n(m)\approx\frac{\mu^m}{m!}e^{-\mu}, \qquad (24)$$

де $\mu = n\cdot p_n$.

За формулою Пуассона обчислюють ймовірність $P_n(m)$ появи події A m разів в n випробуваннях, якщо число випробувань n є досить великим, $P(A)=p_n\to 0$ при $n\to\infty$ і $n\cdot p_n=\mu=const$. Цю формулу можна також використовувати, коли число випробувань n *>200*, а *np <10*.

Приклад 40. При виробництві персональних комп'ютерів ймовірність випуску нестандартного ПК 0,01. Яка ймовірність, що в партії з 200 ПК 2 з них будуть нестандартними?

Розв'язування. Нехай подія А={навмання вибраний ПК є нестандартним}. За умовою задачі ймовірність $p=P(A)=0{,}01$ є малою, число $n=200$ - достатньо велике. Тому для обчислення ймовірності $P_{200}(2)$ доцільно використати формулу Пуассона (24).

В нашому випадку $\mu = n\cdot p = 200\cdot 0{,}01 = 2$ і

$$P_{200}(2)\approx\frac{\mu^2}{2!}e^{-\mu} = \frac{4}{2}e^{-2} = 0{,}2707.$$

§6. *Ймовірність відхилення відносної частоти випадкової події від її ймовірності*

Нехай ймовірність появи випадкової події A в кожному з n незалежних випробувань за схемою Бернуллі дорівнює $p = P(A)$. Відносна частота події A дорівнює $W(A) = \frac{m}{n}$, де m - число появи події A в цих випробуваннях.

Оцінимо ймовірність випадкової події $|W(A) - p| \le \varepsilon,$ де $\varepsilon > 0$ і є малою величиною. Іншими словами, оцінимо ймовірність того, що відхилення відносної частоти $W(A)$ події A від її ймовірності $p = P(A)$ за абсолютним значенням не перевищує заданого малого числа $\varepsilon > 0$.

Нерівність $|W(A) - p| \le \varepsilon, \varepsilon > 0$ перетворимо до рівносильного вигляду

$$|W(A) - p| \le \varepsilon \Leftrightarrow \left|\frac{m}{n} - p\right| \le \varepsilon \Leftrightarrow \left|\frac{m-np}{n}\right| \le \varepsilon \Leftrightarrow -\varepsilon \le \frac{m-np}{n} \le \varepsilon \Leftrightarrow$$
$$\Leftrightarrow -\varepsilon\sqrt{\frac{n}{pq}} \le \frac{m-np}{\sqrt{npq}} \le \varepsilon\sqrt{\frac{n}{pq}} \quad . \tag{25}$$

Оскільки початкова і кінцева нерівності в ланцюжку (25) рівносильні, то

$$P(|W(A) - p| \le \varepsilon) = P\left(-\varepsilon\sqrt{\frac{n}{pq}} \le \frac{m-np}{\sqrt{npq}} \le \varepsilon\frac{n}{pq}\right). \tag{26}$$

З другої сторони, маємо такі нерівності:

$$m_1 \le m \le m_2 \Leftrightarrow m_1 - np \le m - np \le m_2 - np \Leftrightarrow$$
$$\Leftrightarrow \frac{m_1 np}{\sqrt{npq}} \le \frac{m-np}{\sqrt{npq}} \le \frac{m_2 - np}{\sqrt{npq}} \quad . \tag{27}$$

Далі, використовуючи інтегральну формулу Лапласа, одержимо:

$$P\left(\frac{m_1 np}{\sqrt{npq}} \le \frac{m-np}{\sqrt{npq}} \le \frac{m_2 - np}{\sqrt{npq}}\right) = P(m_1 \le m \le m_2) = \Phi(x_2) - \Phi(x_1), \tag{28}$$

де $x_1 = \frac{m_1 - np}{\sqrt{npq}}, x_2 = \frac{m_2 - np}{\sqrt{npq}}$

Рівність (28) зводиться до вигляду:

$$P\left(-\varepsilon\sqrt{\frac{n}{pq}} \le \frac{m-np}{\sqrt{npq}} \le \varepsilon\sqrt{\frac{n}{pq}}\right) = \Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right) \quad , \tag{29}$$

тут $m_1 = np - \varepsilon n, \quad m_2 = np + \varepsilon n$.

Об'єднавши рівності (26) і (29) і врахувавши непарність функції Лапласа $\Phi(x)$, отримуємо розв'язок задачі у такому вигляді:

$$P(|W(A) - p| \le \varepsilon) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right), \tag{30}$$

тобто ***ймовірність того, що при*** n ***незалежних випробування за схемою Бернуллі відхилення відносної частоти*** $W(A)$ ***випадкової події*** A ***від***

ймовірності $p = P(A)$ ***цієї події не перевищує числа*** $\varepsilon > 0$***, дорівнює подвоєному значенню функції Лапласа*** $\Phi(x)$ ***при*** $x = \varepsilon\sqrt{\dfrac{n}{pq}}$ **.**

Приклад 41. Монета підкидається *100* разів. Випадкова подія *A* – поява герба має ймовірність *p=0,5*. Знайти межі, в яких буде міститися відносна частота події *A* з ймовірністю *0,9544*.

Розв'язування. За умовою задачі $n = 100, p = 0{,}5, q = 0{,}5, P(|W(A) - 0{,}5| \le \varepsilon) = 0{,}9544$

За формулою (30) маємо, що

$$0{,}95 = 2\Phi\left(\varepsilon\sqrt{\frac{100}{0{,}25}}\right) \Leftrightarrow 0{,}4772 = \Phi(20\varepsilon)$$

За таблицею значень функції Лапласа $\Phi(x)$ знаходимо, що $20\varepsilon = 2 \Rightarrow \varepsilon = 0{,}1$

З нерівності

$$|W(A) - 0{,}5| < 0{,}1 \Leftrightarrow -0{,}1 \le W(A) - 0{,}5 \le 0{,}1 \Leftrightarrow 0{,}4 \le W(A) \le 0{,}6$$

робимо висновок, що відносна частота $W(A)$ події A міститься в межах від *0,4* до *0,6*.

Враховуючи, що $W(A) = \dfrac{m}{n}$, маємо:

$$0{,}4 \le \frac{m}{100} \le 0{,}6 \Leftrightarrow 40 \le m \le 60\,,$$

тобто у проведених *100* незалежних випробуваннях за схемою Бернуллі подія A з ймовірністю *0,9544* з'явиться від *40* до *60* разів.

Найімовірніше число m_0 появи події A визначається з рівності

$$100 \cdot 0{,}5 - 0{,}5 < m_0 < 100 \cdot 0{,}5 + 0{,}5 \Rightarrow 49{,}5 \le m_0 \le 50{,}5 \Rightarrow m_0 = 50\,.$$

Завдання для самостійної роботи студентів

1. Ймовірність того, що інвестиційний проект принесе через рік прибуток, дорівнює 0,9. Знайдіть імовірності наступних подій. З чотирьох інвестиційних проектів виявляться прибутковими:

а) два; б) всі; в) більше двох; г) хоча б один.

а) Ймовірність того, що відвідувач магазину зробить покупку, дорівнює 0,4. Знайдіть імовірності наступних подій. Із 10 відвідувачів, які завітали до магазину, покупку зроблять: а) 3 відвідувача; б) від 2-х до 4-х; в) більше4-х;

2. На біржі виставлено 20 цінних паперів. Імовірність того, що вони подорожчають протягом дня, дорівнює 0,7. Знайдіть імовірності наступних подій:

а) подорожчають 6 цінних паперів;
б) жоден з цінних паперів не подорожчає;
в) подорожчають не більше 3-х цінних паперів;
г) подорожчають більше 9-и цінних паперів;
д) всі цінні папери подорожчають.

3. Інвестор укладає договір на фондовій біржі. Ймовірність укладання однієї угоди за один день складає 0,7. Виходячи з припущення, що за період у 20 робочих днів щодня укладається не більше однієї угоди, знайти ймовірності наступних подій:

а) будуть укладені рівно 4 угоди;

б) будуть укладені 12 угод;

в) буде укладена хоча б одна угода;

г) будуть укладені не менше 7-и угод;

д) будуть укладені менше 4-х угод.

4. У виробництві певної продукції третій сорт складає 10 %. Знайдіть імовірність того, що з п'яти навмання взятих виробів цієї продукції не менш трьох будуть третього сорту. Яке найімовірніше число виробів третього сорту серед 4-х відібраних?

5. Для забезпечення нормальної роботи відділення банку потрібно, щоб діючими були не менше, ніж 90% із наявних 85-ти комп'ютерів. Яка ймовірність нормальної роботи відділення банку, якщо ймовірність вийти з ладу для кожного комп'ютера складає 0,06?

6. Ймовірність вчасної реалізації одиниці продукції складає 0,75. Визначити ймовірність вчасної реалізації не менше 425-ти одиниць продукції із 500, що поступили у реалізацію.

7. Ймовірність влучення в "десятку" при одному пострілі складає 0,4. Скільки пострілів слід здійснити, щоб з ймовірністю не менше, ніж 0,85 влучити в "десятку" хоча б один раз?

8. На складі є продукція двох сортів, причому виробів другого сорту в 2 рази більше виробів першого сорту. Знайдіть імовірність того, що серед чотирьох навмання взятих виробів хоча б один першого сорту. Яке найімовірніше число виробів першого сорту з чотирьох відібраних?

9. Ймовірність виготовлення нестандартної деталі дорівнює 0,07. Скільки треба виготовити деталей, щоб найімовірніше число нестандартних деталей було 65?

10. Скільки разів треба кинути гральний кубик, щоб найімовірніше число появи трійки дорівнювало 50?

11. Ймовірність народження хлопчика дорівнює 0,5. Скільки треба мати дітей в сім'ї, щоб з імовірністю не менше 0,95 там був хоча б один хлопчик?

12. Ймовірність хоча б одного успіху при чотирьох незалежних випробуваннях дорівнює 0,52. Яка ймовірність успіху при одному випробуванні?

13. Ймовірність того, що банк отримає денну норму грошей, необхідну для функціонування, складає 0,94. Знайдіть імовірність того, що із 200 робочих днів на протязі:

а) 170 днів денна норма отриманих коштів буде виконана;

б) число днів, коли буде виконана норма отримання необхідних грошей, знаходиться в межах від 170 до 185 днів.

14. Скільки треба провести випробувань з киданням монети, щоб з імовірністю не менше 0,95 можна було сподіватися, що відхилення відносної частоти випадіння герба буде відрізнятися за абсолютною величиною від імовірності менш ніж 0,5?

15. У страховій компанії зареєстровано 2000 автомобілів. Імовірність поломки будь-якого автомобіля під час аварії дорівнює 0,005. Кожен власник, що застрахував свій автомобіль, сплачує 200 грн за рік, а у випадку поломки отримує 1000 грн. Знайдіть імовірність того, що після року праці компанія зазнає збитку.

16. Два рівносильні гравці грають в шахи (нічиї до уваги не приймаються). Що ймовірніше:

а) виграти дві партії з чотирьох чи три партії з шести;

б) виграти одну партію з двох чи дві партії з чотирьох;

в) виграти не менше двох партій з чотирьох чи не менше трьох партій з п’яти?

17. В пологовому будинку народилося 10 немовлят. Відомо, що ймовірність народження хлопчика дорівнює 0,51. Знайти ймовірність, що серед цих немовлят:

а) два хлопчики;

б) не більше двох хлопчиків;

в) більше, ніж два хлопчики;

г) Яка найімовірніша кількість хлопчиків народиться?

18. Ймовірність попадання в ціль при одному пострілі дорівнює 0,7 Знайти ймовірність того, що при 200 пострілах попадання в ціль буде рівно 145 разів.

19. Фірма, що проводить опитування, очікує, що 60% опитуваних заповнять анкети. Опитано 30 сімей. Яка ймовірність того, що заповнять анкети а) 12 сімей; б) 25 сімей?

20. Ймовірність несплати податку для кожного із 100 підприємств дорівнює 0,2. Яка ймовірність, що податки не сплатять не більше 17-ти підприємств?

21. У відділенні ПІБу працює 450 співробітників. Знайти ймовірність того, що у двох співробітників день народження припаде на 1-е травня. Вважати, що кількість днів у році 365.

22. Підручник видано тиражем 500 екземплярів. Ймовірність того, що підручник зверстано неправильно, дорівнює 0,01. Знайти ймовірність, що тираж містить рівно п’ять бракованих книг.

23. Магазин одержав 500 пляшок мінеральної води. Ймовірність того, що при транспортуванні пляшка розіб’ється, дорівнює 0,007. Знайти ймовірність, що магазин отримає розбитих пляшок: а) рівно три; б) менше трьох; в) більше ніж три; г) хоча б одну.

24. Середня кількість викликів, які надходять в диспетчерський пункт ЖЕКу за годину дорівнює шість. Знайти ймовірність, що за дві години надійде: а) чотири виклики; б) менше чотирьох викликів; в) не менше шести викликів.

25. Середня кількість викликів, які надходять на АТС за одну секунду дорівнює два. Знайти ймовірність, що за чотири секунди надійде: а) три виклики; б) менше трьох викликів; в) не менше трьох викликів.

Розділ 2. Випадкові величини. Функції розподілу

Глава 1. Випадкові величини

§ 1. Поняття випадкової величини

Одним з основних понять теорії ймовірностей є поняття випадкової величини,з яким пов'язане уявлення про стохастичний експеримент, що полягає у вимірюванні певної числової величини $\boldsymbol{X}$. Прикладами випадкових величин можуть бути кількість очок, що випадають при одному підкиданні грального кубика, кількість влучень у ціль при $\boldsymbol{n}$ пострілах, розмір виграшу на лотерейний білет під час розіграшу лотереї, час безвідмовної роботи приладу, показники фінансових, економічних процесів, кількість відвідувачів у день відділення банку та інші.

Означення. ***Випадковою величиною X називається числова функція, яка визначена на просторі елементарних подій Ω={ω}, така що для кожного дійсного a існує ймовірність P(X<a)= P(ω:X(ω)<a).***

Надалі випадкові величини позначатимемо великими літерами латинського алфавіту $X, Y, Z, ...$, а їх можливі значення – відповідно малими літерами $x, y, z, ...$

Означення. ***Випадкова величина називається дискретною, якщо вона набуває окремих, ізольованих можливих значень з певними ймовірностями.***

Множина можливих значень дискретної випадкової величини може бути скінченною або зліченною.

Означення. ***Випадкова величина називається неперервною, якщо вона може набувати всіх значень з деякого скінченного або нескінченного проміжку.***

§ 2. *Дискретні випадкові величини. Закони розподілу дискретної випадкової величини*

Нехай X - дискретна випадкова величина, можливими значеннями якої є числа $x_1, x_2, ... x_n$. Через $p_i = P(X = x_i)$ позначимо ймовірності цих значень, тобто p_i є ймовірність випадкової події, яка полягає в тому, що випадкова величина X набуває значення x_i. Події $X - x_i \left(i = \overline{1, n}\right)$ утворюють повну групу подій, тому

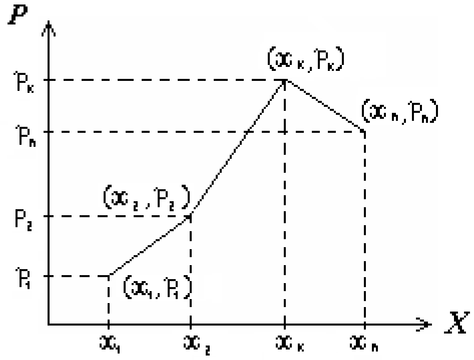
$$p_1 + p_2 + ... + p_n = 1.$$

Означення. ***Законом розподілу дискретної випадкової величини називається відповідність між всіма можливими значеннями цієї величини та їх ймовірностями.***

Закон розподілу дискретної випадкової величини подають у вигляді таблиці значень випадкової величини і відповідних ймовірностей.

Для графічного зображення закону розподілу дискретної випадкової величини на прямокутну систему координат наносять точки (x_i, p_i) послідовно з'єднують їх відрізками. Одержану фігуру називають ***многокутником розподілу.***

Для цього обираємо систему координат. На вісі абсцис відкладаємо можливі значення дискретної випадкової величини, а на вісі ординат – відповідні значення ймовірностей **P.** При цьому знаходимо точки з координатами (x_1, p_1) , (x_2, p_2) , ... , (x_n, p_n). З'єднуючи послідовно одержані точки відрізками прямих, дістанемо графік (мал. 5) у вигляді многокутника розподілу дискретної випадкової величини.



Мал. 5

Приклад 42. Робітник обслуговує три незалежно працюючих верстати. Ймовірність того, що на протязі години верстату знадобиться увага робітника, дорівнює для першого верстата 0,3, для другого – 0,2, для третього – 0,1. Знайдіть а) закон розподілу випадкової величини X – числа верстатів, котрим знадобиться увага робітника на протязі години, б) многокутник розподілу.

Розв'язування. Нехай подія A – першому верстату на протязі години знадобиться увага робітника, подія B – другому знадобиться увага робітника, C – третьому. За умовою задачі $P(A)=0{,}3$, $P(B)=0{,}2$, $P(C)=0{,}1$. Випадкова величина X – число верстатів, котрим знадобиться увага робітника на протязі години, приймає значення *0, 1, 2, 3*. Очевидно, що

$$P\{X=0\}=P\left(\overline{A}\cap\overline{B}\cap\overline{C}\right)=P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\cdot P\left(\overline{C}\right)=0{,}7\cdot 0{,}8\cdot 09=0{,}504\,;$$

$$P\{X=1\}=P\left(A\cap\overline{B}\cap\overline{C}\right)+P\left(\overline{A}\cap B\cap\overline{C}\right)+P\left(\overline{A}\cap\overline{B}\cap C\right)=$$

$$=0{,}3\cdot 0{,}8\cdot 0{,}9+0{,}7\cdot 0{,}2\cdot 0{,}9+0{,}7\cdot 0{,}8\cdot 0{,}1=0{,}398;$$

$$P\{X=2\}=P\left(\overline{A}\cap B\cap C\right)+P\left(A\cap\overline{B}\cap C\right)+P\left(A\cap B\cap\overline{C}\right)=$$

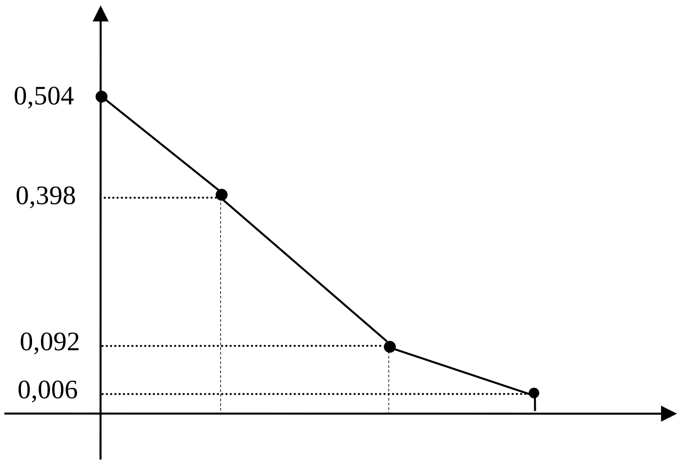
$$=0{,}7\cdot 0{,}2\cdot 0{,}1+0{,}3\cdot 0{,}8\cdot 0{,}1+0{,}3\cdot 0{,}2\cdot 0{,}9=0{,}092;$$

$$P\{X=3\}=P(A\cap B\cap C)=P(A)\cdot P(B)\cdot P(C)=0{,}3\cdot 0{,}2\cdot 0{,}1=0{,}006\,.$$

а) закон розподілу випадкової величини X має вигляд:

X	0	1	2	3
p	0,504	0,398	0,092	0,006

б) многокутник розподілу має вигляд:



Мал. 6

Приклад 43. В грошовій лотереї розігрується 2 виграші по 1000 грн., 10 виграшів по 100 грн., і 100 виграшів по 10 грн. При загальному числі лотерейних квитків 10000. Знайти закон розподілу випадкової величини X $x_1 = 1000, x_2 = 100, x_3 = 10, x_4 = 0$.

Розв'язування. Ймовірності цих значень обчислимо за формулою класичної ймовірності:

$$p_1 = \frac{2}{10000} = 0{,}0002;\ p_2 = \frac{10}{10000} = 0{,}001;\ p_3 = \frac{100}{10000} = 0{,}01;$$

$$p_4 = 1 - (p_1 + p_2 + p_3) = 1 - (0{,}0002 + 0{,}001 + 0{,}01) = 0{,}9888.$$

Закон розподілу випадкової величини X запишемо у вигляді таблиці:

X	1000	100	10	0
P	0.0002	0.001	0.01	0.9888

Наведемо найбільш важливі дискретні розподіли випадкових величин.

Біноміальний розподіл

Нехай проводиться n незалежних випробувань, в кожному з яких випадкова подія може з'явитися (успіх) або не з'явитися (невдача). Ймовірність появи успіху у кожному випробуванні є сталою і дорівнює p (ймовірність не появи успіху дорівнює $q = 1 - p$). Випадкова величина X - число появ успіхів в цих випробуваннях. Знайдемо закон розподілу дискретної випадкової величини X. Величина X може набувати значень $0,1,2,3,...,n$. Для обчислення ймовірностей можливих значень випадкової величини X використаємо формулу Бернуллі

$$\overline{P_i} = P_n(i) = C_n^i p^i q^{n-i} \tag{31}$$

і дістанемо закон розподілу випадкової величини X у вигляді таблиці

X	0	1	...	$\boldsymbol{n}$
p	$P_n(0)=C_n^0 p^0 q^n$	$P_n(1)=C_n^1 p^1 q^{n-1}$	...	$P_n(n)=C_n^n p^n q^0$

Одержаний закон розподілу дискретної випадкової величини X у вигляді таблиці називається ***біноміальним.*** Бачимо, що бернулієва випадкова величина визначається двома параметрами $\boldsymbol{n}$ та $\boldsymbol{p}$.

Приклад 44. Прилад складається з 4-х елементів і ймовірність несправності кожного з них дорівнює $\frac{1}{8}$. Скласти закон розподілу числа несправних елементів приладу. Визначити ймовірність того, що число несправних елементів більше від двох.

Розв’язування. Випадкова величина X - число несправних елементів приладу може набувати значень $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3, x_5 = 4.$ Ймовірність цих можливих значень обчислимо за формулою Бернуллі:

$$p_1 = C_4^0 p^0 q^4 = 1\cdot 1\cdot\left(\frac{1}{2}\right)^4 = \frac{1}{16};$$

$$p_2 = C_4^1 p^1 q^3 = 4\cdot\frac{1}{2}\cdot\left(\frac{1}{2}\right)^3 = \frac{1}{4};$$

$$p_3 = C_4^2 p^2 q^2 = \frac{4!}{2!2!}\cdot\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)^2 = \frac{3}{8};$$

$$p_4 = C_4^3\left(\frac{1}{2}\right)^3\cdot\frac{1}{2} = \frac{1}{4};$$

$$p_5 = 1-\left(\frac{1}{16}+\frac{1}{4}+\frac{3}{8}+\frac{1}{4}\right) = \frac{1}{16}.$$

Закон розподілу величини X має вигляд:

X	0	1	2	3	4
p	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Ймовірність події $B = \{$число несправних елементів більше, ніж два$\}$ визначається з одержаної таблиці:

$$P(B) = \frac{1}{4} + \frac{1}{16} = \frac{5}{16}.$$

З таблиці також видно, що найбільш ймовірне число несправних елементів у приладі є $m_0 = 2$.

Розподіл Пуассона

Нехай величина X набуває значень $0,1,2,3,...,n,...$, а ймовірність появи відповідного значення обчислюється за формулою:

$$p_i = \frac{(\mu)^i}{i!} e^{-\mu} \tag{32}$$

Такий закон розподілу дискретної випадкової величини X називають ***розподілом Пуассона:***

X	0	1	...	n	...
p	$p_0 = \frac{(\mu)^0}{0!} e^{-\mu}$	$p_1 = \frac{(\mu)^1}{1!} e^{-\mu}$	...	$p_n = \frac{(\mu)^n}{n!} e^{-\mu}$	...

Бачимо, що пуассонівська випадкова величина визначається одним параметром μ.

Приклад 45. Електронна пошта банку підтримує зв'язок із сотнею абонентів. Ймовірність того, що за одиницю часу на електронну пошту банку надійде повідомлення, дорівнює $p = 0{,}02$. Записати закон розподілу числа сигналів від 4-х абонентів. Яка з подій є більш ймовірною: за одиницю часу поступлять сигнали від 3-х чи 4-х абонентів?

Розв'язування. Нехай випадкова величина X - число сигналів, що поступлять на електронну пошту від 4-х абонентів. Величина X може набути значення $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3, x_5 = 4.$ Оскільки ймовірність p появи події $A =$ {поступлення сигналу на електронну пошту} є мала, то для обчислення ймовірностей можливих значень величини X використаємо формулу Пуассона (32).

В нашому випадку $n = 100, p = 0{,}02, \mu = 100 \cdot 0{,}02 = 2$. Обчислимо ймовірності p_i можливих значень величини X:

$$p_1 = P_{100}(0) = e^{-2} = 0{,}1353; \quad p_2 = P_{100}(1) = 2e^{-2} = 0{,}2706;$$

$$p_3 = P_{100}(2) = 2e^{-2} = 0{,}2706; \quad p_4 = P_{100}(3) = \frac{4}{3}e^{-2} = 0{,}1804;$$

$$p_5 = 1 - (0{,}1353 + 0{,}2706 + 0{,}2706 + 0{,}1804) = 1 - 0{,}3569 = 0{,}1431;$$

Шуканий закон розподілу має вигляд:

X	0	1	2	3	4
p	0,1353	0,2706	0,2706	0,1804	0,1431

Більше ймовірно, що сигнали поступлять від 3-х абонентів, ніж від 4-х, оскільки: $p_4 = P_{100}(3) = 0{,}1804$, а $p_5 = P_{100}(4) = 0{,}1431$.

Найбільш ймовірне число сигналів, що поступлять на електронну пошту, є один або два.

Геометричний розподіл

Нехай проводиться n незалежних випробувань, ймовірність появи випадкової події A в кожному з них дорівнює p $(0 < p < 1)$, а ймовірність не появи цієї події дорівнює $q = 1 - p$. Випробування закінчуються, як тільки з'явиться подія A. Це означає, що коли подія A з'явилася в $m-$му випробуванні, то в попередніх $m-1$ випробуваннях вона не з'явилася.

Позначимо через X дискретну випадкову величину – число випробувань, які потрібно провести до першої появи події A. Можливі значення величини X є натуральні числа: $x_1 = 1, x_2 = 2, ...$

Нехай в перших $m-1$ випробуваннях подія A не з'явилася, а в m - тому випробуванні з'явилася. Ймовірність $P(X = m)$ цієї складної події обчислимо за теоремою множення ймовірностей незалежних подій:

$$P(X = m) = q^{m-1} p \tag{33}$$

(q^{m-1} - ймовірність того, що подія A в перших $m-1$ випробуваннях не з'явилася, p - ймовірність того, що в m - му випробуванні подія A - з'явилася).

Якщо у формулі (33) надавати значення $m = 1,2,3,...,$ то одержимо послідовність

$$p, qp, q^2 p, ..., q^{m-1} p, ..., \tag{34}$$

яка є геометричною прогресією з першим членом p і знаменником q. Тому закон розподілу дискретної випадкової величини X, що виражається формулою (33) називається ***геометричним.***

Легко переконатися, що послідовність (34) збігається $(q < 1)$ і сума її членів

$$p + qp + q^2 p + ... + q^{m-1} p + ... = \frac{p}{1-q} = 1$$

Закон розподілу описаної дискретної випадкової величини X можна записати у вигляді таблиці:

X	1	2	...	m	...
p	$P(X = 1) = p$	$P(X = 2) = q\ p$	...	$P(X = m) = q^{m-1} p$	...

Приклад 46. З мінометa стріляють по цілі до першого влучення. Ймовірність влучення в ціль $p = 0{,}6$. Знайти закон розподілу випадкової величини ***X*** - кількість пострілів, і ймовірність того, що влучення настане при третьому пострілі.

Розв'язування. Використаємо формулу (33) і запишемо закон розподілу:

X	*1*	*2*	*3*	...	*к*	...
p	*0,6*	*0,4.0,6*	*0,4*^{2}*.0,6*		*0,4*^{к}*.0,6*	...

Ймовірність влучення при третьому пострілі дорівнює:

$$P(X = 3) = 0{,}4^2 \cdot 0{,}6 = 0{,}16 \cdot 0{,}6 = 0{,}096.$$

Гіпергеометричний розподіл

Приклад 47. В урні знаходяться $\boldsymbol{N}$ куль серед яких $\boldsymbol{n}$ білих і $\boldsymbol{N-n}$ чорних. Навмання з урни витягли k куль. Нехай $\boldsymbol{X}$ – кількість білих серед k відібраних куль. Знайти закон розподілу дискретної випадкової величини $\boldsymbol{X}$.

Розв'язування. Очевидно, число всіх партій з k куль дорівнює C_N^k, а серед них $C_N^r \cdot C_{N-n}^{k-r}$ партій, що містять r білих куль. Отже, за класичним означенням ймовірностей маємо :

$$P(X=r)=\frac{C_N^r C_{N-n}^{k-r}}{C_N^k} \quad (0 \le r \le \min(n,k)). \qquad (35)$$

Ряд ймовірностей, що знаходяться за формулою (35), називається гіпергеометричним розподілом.

Приклад 48. У лотереї "Спортлото" вказано 49 видів спорту. Студент на картці відмітив шість видів спорту. Після того, як він відправив картку, проведено тираж (із $\boldsymbol{N}$ по k).

Якщо виявиться, що не менше $s=3$ з них збігається з тими, що занесені на картку, то студент отримує виграш. Знайти ймовірність виграшу.

Розв'язування. Нехай A_r - подія , яка означає, що студент вгадав r видів спорту. Тоді за формулою (35) ймовірність події буде:

$$P(A_r)=\frac{C_k^r \cdot C_{N-k}^{k-r}}{C_N^k},$$

а ймовірність виграшу (подія $\boldsymbol{A}$) складе :

$$P(A)=\sum_{r=s}^{k} P(A_r)=\frac{\sum_{r=s}^{k} C_k^r \cdot C_{N-k}^{k-r}}{C_N^k}..$$

Закон розподілу дискретної випадкової величини $\boldsymbol{X}$ матиме вигляд :

X	*0*	*1*	*2*	*3*	*4*	*5*	*6*
p	*0,435965*	*0,413019*	*0,132378*	*0,017650*	*0,000969*	*0,000018*	*0,000000*

Отже, шукана ймовірність виграшу дорівнює:

$$P(A)=P(A_3)+P(A_4)+P(A_5)+P(A_6)=0{,}018638.$$

§3. Функція розподілу випадкової величини

Означення. Якщо для довільного дійсного числа x визначена ймовірність того, що X приймає значення менші за x, то ця ймовірність:

$$P\{X < x\} = F(x)$$

називається ***функцією розподілу*** випадкової величини X.

Властивості функції розподілу:

1) $0 \le F(x) \le 1$;
2) $F(-\infty) = 0$, $F(+\infty) = 1$;
3) $F(x_1) \le F(x_2)$, якщо $x_1 < x_2$;
4) $P\{a \le X < b\} = F(b) - F(a)$;
5) функція розподілу неперервна зліва, тобто $\lim\limits_{x \to x_0 - 0} F(x) = F(x_0)$.

Функція розподілу дискретної випадкової величини:

$$F(x) = P\{X < x\} = \sum_{x_i < x} p_i \tag{36}$$

дорівнює сумі тих ймовірностей, при яких значення випадкової величини менші за x.

Приклад 49. В умовах прикладу 42 знайдемо функцію розподілу випадкової величини X – числа верстатів, котрим знадобиться увага робітника протягом години:

Розв’язування. Закон розподілу випадкової величини X має вигляд:

X	*0*	*1*	*2*	*3*
p	*0,504*	*0,398*	*0,092*	*0,006*

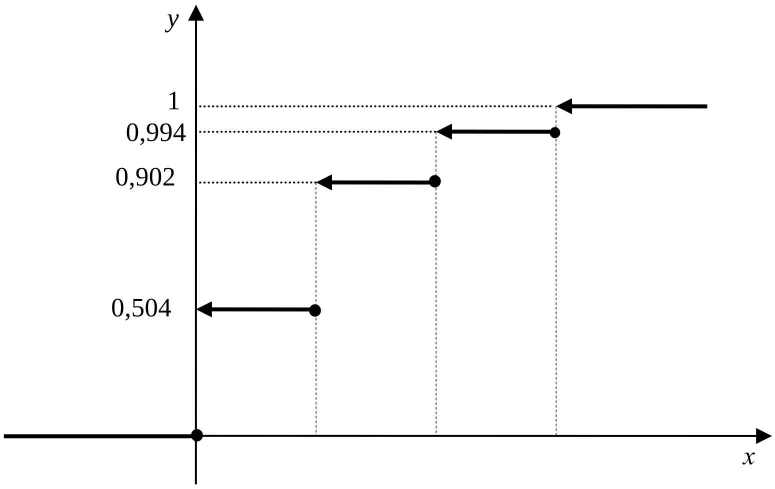
Побудуємо функцію розподілу випадкової величини X:

1. при $x \le 0$ $F(x) = \sum\limits_{x_i < x} p_i = 0$;
2. при $0 < x \le 1$ $F(x) = \sum\limits_{x_i < x} p_i = P(X = 0) = 0{,}504$;
3. при $1 < x \le 2$ $F(x) = \sum\limits_{x_i < x} p_i = P(X = 0) + P(X = 1) = 0{,}504 + 0{,}398 = 0{,}902$;
4. при $2 < x \le 3$ $F(x) = \sum\limits_{x_i < x} p_i = P(X = 0) + P(X = 1) + P(X = 2) = 0{,}504 + 0{,}398 + 0{,}092 =$ $= 0{,}994$;
5. при $x > 3$ $F(x) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0{,}504 + 0{,}398 + 0{,}092 +$ $+ 0{,}006 = 1$.

Отже, функція розподілу даної випадкової величини X має вигляд:

$$F(x)=\begin{cases}0, & \text{при } x\le 0;\\ 0{,}504, & \text{при } 0<x\le 1;\\ 0{,}902, & \text{при } 1<x\le 2;\\ 0{,}994, & \text{при } 2<x\le 3;\\ 1, & \text{при } x>3\end{cases}.$$

Будуємо графік функції розподілу даної дискретної випадкової величини *X*:



Мал. 7

§4. Неперервні випадкові величини

Означення. Випадкова величина називається ***неперервною***, якщо її функція розподілу може бути представлена у вигляді:

$$F(x)=\int_{-\infty}^{x} f(t)dt, \qquad (37)$$

де функція $f(x)$ називається ***щільністю розподілу ймовірностей*** випадкової величини.

Властивості щільності розподілу:

1) $f(x)=F'(x)$;
2) $f(x)\ge 0$;
3) $\int_{-\infty}^{+\infty} f(x)dx=1$;
4) $P\{a\le X<b\}=\int_{a}^{b} f(x)dx$.

Приклад 50. Задано випадкову величину з щільністю розподілу ймовірностей:

$$f(x)=\begin{cases}A\cos x, & \text{при } -\pi/2\le x\le \pi/2\\ 0, & \text{при } |x|>\pi/2\end{cases}.$$

Знайдемо: а) величину сталої A; б) функцію розподілу $F(x)$; в) $P\left\{0 \le X < \pi/6\right\}$.

Розв'язування. а) для знаходження значення сталої A скористаємось властивістю щільності 3):

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

Тому

$$1 = \int_{-\infty}^{+\infty} f(x)dx = \int_{-\pi/2}^{\pi/2} A\cos x dx = A\sin x\Big|_{-\pi/2}^{\pi/2} = A\left(\sin\frac{\pi}{2} - \sin\left(-\frac{\pi}{2}\right)\right) = 2A;\ A = 1/2.$$

Таким чином, щільність розподілу ймовірностей приймає вигляд:

$$f(x) = \begin{cases} 1/2 \cos x, \text{при} -\pi/2 \le x \le \pi/2 \\ 0, \text{при} |x| > \pi/2 \end{cases}.$$

б) з (37) відомо, що $F(x) = \int_{-\infty}^{x} f(t)dt$. *Знаходимо значення функції розподілу на кожному інтервалі окремо:*

1. $x \le -\pi/2$, $\quad F(x) = \int_{-\infty}^{x} 0dt = 0$;
2. $-\pi/2 < x \le \pi/2$,

$$F(x) = \int_{-\infty}^{-\pi/2} 0dt + \int_{-\pi/2}^{x} \frac{1}{2}\cos t dt = \frac{1}{2}\sin t\Big|_{-\pi/2}^{x} = \frac{1}{2}\left(\sin x - \sin(-\pi/2)\right) = \frac{1}{2}(\sin x + 1);$$

3. $x > \pi/2$, $F(x) = \int_{-\infty}^{-\pi/2} 0dt + \int_{-\pi/2}^{\pi/2} \frac{1}{2}\cos t dt + \int_{\pi/2}^{x} 0dt = \frac{1}{2}\sin t\Big|_{-\pi/2}^{\pi/2} = \frac{1}{2}\left(\sin\pi/2 - \sin\left(-\pi/2\right)\right) = 1.$

Отже, $F(x) = \begin{cases} 0, & x \le -\pi/2 \\ \frac{1}{2}(\sin x + 1), & -\pi/2 < x \le \pi/2 \\ 1, & x > \pi/2 \end{cases}.$

в) використовуючи властивості функції розподілу, знаходимо:

$$P\left\{0 \le X < \pi/6\right\} = F\left(\pi/6\right) - F(0) = \frac{1}{2}\left(\sin\left(\pi/6\right) + 1\right) - \frac{1}{2}\left(\sin 0 + 1\right) = \frac{1}{4}.$$

§5. Математичне сподівання дискретних випадкових величин

Значення, які приймає випадкова величина заздалегідь передбачити не можна. Тому випадкову величину прагнуть характеризувати певними числовими параметрами, до яких належать математичне сподівання, дисперсія, середнє квадратичне відхилення та інші.

Означення. Математичним сподіванням ***дискретної випадкової величини*** X називається число, яке дорівнює сумі добутків усіх її можливих значень на відповідні ймовірності і позначається $M(X)$:

$$M(X)=\sum_{k=1}^{n} x_k p_k. \qquad (38)$$

Тут X – дискретна випадкова величина, що набуває значення x_k з ймовірностями p_k, $k=\overline{1,n}$.

Зауваження 1. Якщо дискретна випадкова величина X набуває зліченну множину можливих значень, то

$$M(X)=\sum_{k=1}^{\infty} x_k p_k, \qquad (39)$$

За умови, що даний ряд абсолютно збіжний. Якщо ряд (39) розбігається , то кажуть, що випадкова величина X не має математичного сподівання.

Зауваження 2. З означення випливає, що $M(X)$ не випадкова, а стала величина (на відміну від випадкової величини X).

Приклад 51. Знайти $M(X)$ випадкової величини X , яка означає кількість появи гербів при підкиданні трьох монет .

Розв'язування. За умовою випадкова величина X означає кількість гербів. Наведемо таблицю усіх можливих значень, яких набуває X , та їх ймовірності.

Ω	*ГГГ*	*ГГЦ*	*ГЦГ*	*ЦГГ*	*ГЦЦ*	*ЦГЦ*	*ЦЦГ*	*ЦЦЦ*
X	*3*	*2*	*2*	*2*	*1*	*1*	*1*	*0*
p	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Тепер запишемо закон розподілу випадкової величини X :

X	*0*	*1*	*2*	*3*
p	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Знайдемо $M(X)$ за формулою (38).

$$M(X)=0\times\frac{1}{8}+1\times\frac{3}{8}+2\times\frac{3}{8}+3\times\frac{1}{8}=1{,}5.$$

Зауваження 3. Математичне сподівання числа появи події A в одному випробуванні дорівнює ймовірності цієї події . Дійсно, $M(X)=0\cdot q+1\cdot p=p.$

Зауваження 4. Математичне сподівання наближено дорівнює (тим точніше, чим більше число випробувань *n*) середньому арифметичному значень випадкової величини, що спостерігається. У цьому полягає ймовірнісний зміст $M(X)$.

Зауваження 5. Очевидно, $M(X)$ більше найменшого і менше найбільшого з можливих значень дискретної випадкової величини *X*:

$$\min(x_1, x_2, ..., x_n) < M(X) < \max(x_1, x_2, ..., x_n).$$

Властивості математичного сподівання дискретних випадкових величин

1. Математичне сподівання сталої величини дорівнює цій сталій

$$M(C) = C. \qquad (40)$$

2. Сталий множник можна виносити за знак математичного сподівання

$$M(C) = C \cdot M(X). \qquad (41)$$

Наведені властивості випливають із означення математичного сподівання.

3. Математичне сподівання добутку двох незалежних дискретних випадкових величин дорівнює добутку їх математичних сподівань

$$M(X \cdot Y) = M(X) \cdot M(Y). \qquad (42)$$

Наслідок. Математичне сподівання добутку декількох взаємно незалежних дискретних випадкових величин дорівнює добутку їх математичних сподівань.

Приклад 52. Незалежні випадкові величини X та Y задані законами розподілу :

X	*2*	*4*	*5*
p	*0,1*	*0,3*	*0,6*

X	*7*	*9*
p	*0,8*	*0,2*

Знайти математичне сподівання $M(XY)$.

Розв'язування. Знайдемо $M(X) = 0{,}2 + 1{,}2 + 3 = 4{,}4;$ $M(Y) = 5{,}6 + 1{,}8 = 7{,}4.$

З рівності (42) знаходимо $M(XY) = 4{,}4 \cdot 7{,}4 = 32{,}56.$

4. Математичне сподівання суми двох дискретних випадкових величин дорівнює сумі математичних сподівань доданків :

$$M(X + Y) = M(X) + M(Y). \qquad (43)$$

Наслідок. Математичне сподівання суми декількох випадкових величин дорівнює сумі математичних сподівань доданків.

Приклад 53. Проведено три постріли з ймовірностями влучення в ціль $p_1 = 0{,}4;\ p_2 = 0{,}3;\ p_3 = 0{,}6.$ Знайти математичне сподівання загальної кількості влучень.

Розв'язування. Кількість влучень при першому пострілі є випадковою величиною , яка може набувати двох значень : 1 (влучення) з ймовірністю $p = 0{,}4$ і 0 (промах) з ймовірністю $q = 1 - p = 1 - 0{,}4 = 0{,}6.$

Тоді, $M(X) = 1 \cdot 0{,}4 + 0 \cdot 0{,}6 = 0{,}4.$ Аналогічно для другого і третього пострілів $M(Y) = 1 \cdot 0{,}3 + 0 \cdot 0{,}7 = 0{,}3$ і $M(Z) = 0{,}6.$

Загальна кількість влучень є випадковою величиною $X + Y + Z,$ тому

$$M(X + Y + Z) = M(X) + M(Y) + M(Z) = 0{,}4 + 0{,}3 + 0{,}6 = 1{,}3.$$

Відхилення випадкової величини від її математичного сподівання

Математичне сподівання не дає повної характеристики випадкової величини, Розглянемо приклад.

Приклад 54. Дві випадкової величини X та Y задано розподілами:

X	*-0,01*	*0,01*
p	*0,5*	*0,5*

Y	*-100*	*100*
p	*0,5*	*0,5*

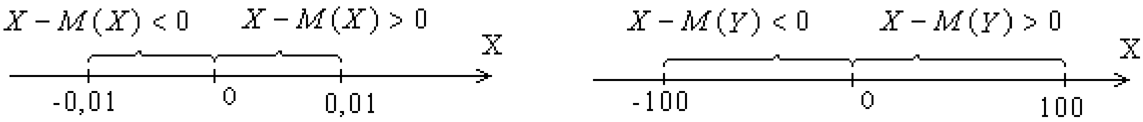


Рис. 8

Знайдемо їхні математичні сподівання $M(X)$ і $M(Y)$:

$$M(X) = -\frac{1}{2} \cdot 0{,}01 + \frac{1}{2} \cdot 0{,}01 = 0, \quad M(Y) = -\frac{1}{2} \cdot 100 + \frac{1}{2} \cdot 100 = 0.$$

Отримали, що математичні сподівання $M(X)$ і $M(Y)$ однакові, хоча поведінка випадкових величин X та Y суттєво відрізняються одне від одного.

Тому вводять нову величину, що характеризує відхилення випадкової величини X від $M(X)$.

Означення. ***Відхиленням*** випадкової величини X називається різниця між випадковою величиною та її математичним сподіванням, тобто $X - M(X)$.

Покажемо, що математичне сподівання відхилення випадкової величини від свого математичного сподівання дорівнює нулю: $M(X - M(X)) = 0.$

Дійсно , згідно з властивостями $M(X)$ маємо :

$$M(X - M(X)) = M(X) - M(M(X)) = M(X) - M(X) = 0.$$

Отже, введена величина теж недостатньо характеризує розташування значень випадкової величини. Тому в теорії ймовірностей розглядається квадрат відхилення значень випадкової величини від її математичного сподівання.

Обчислення математичного сподівання дискретних випадкових величин

Біномний розподіл

Доведемо, що математичне сподівання $M(X)$ числа появи події A (успіхів) в n незалежних випробуваннях дорівнює добутку числа випробувань n на ймовірність p $(0 < p < 1)$ появи події A в кожному випробуванні:

$$M(X) = np.$$

Дійсно.

$$M(X) = \sum_{k=0}^{n} k P_n(k) = \sum_{k=0}^{n} k C_n^k p^k q^{n-k} = np \sum_{k=0}^{n} \frac{k}{n} C_n^k p^{k-1} q^{n-k} =$$

$$= np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{n-k} = np \sum_{r=0}^{n-1} C_{n-1}^{r} p^{k-1} q^{n-1-r} = np.$$

тут враховано, що $\left(\sum_{r=0}^{n-1} C_{n-1}^{r} p^r q^{n-1-r} = 1 \right)$.

Приклад 55. Ймовірність влучення в ціль при стрільбі по мішені дорівнює $p = 0{,}6$. Знайти математичне сподівання $M(X)$ загального числа влучень, якщо зроблено 20 пострілів.

Розв'язування. $M(X) = np = 20 \cdot 0{,}6 = 12$ (влучень).

Геометричний розподіл дискретної випадкової величини

Покажемо, що якщо дискретна випадкова величина X має геометричний розподіл, то її математичне сподівання $M(X) = 1/p$, де $p(0 < p < 1)$ – ймовірність появи події в одному випробуванні.

Дійсно. За припущенням дискретна випадкова величина X має геометричний розподіл. Тому:

$$P(X = k) = q^{k-1} p, \left(k = \overline{1, \infty}\right).$$

Тоді,

$$M(X) = \sum_{k=1}^{\infty} k q^{k-1} p = p \cdot (1 + 2q + 3q^2 + \cdots + nq^{n-1} + \cdots) = p \cdot \frac{1}{(1-q)} = p \cdot \frac{1}{p^2} = \frac{1}{p}.$$

Розподіл Пуассона

Покажемо, що якщо дискретна випадкова величина X має розподіл Пуассона з параметром λ, то її математичне сподівання $M(X) = \lambda.$.

Дійсно. За умовою дискретна випадкова величина має розподіл Пуассона з параметром λ, тому $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, (k = \overline{0, \infty}).$

Отже,

$$M(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda..$$

Приклад 56. Дискретна випадкова величина задана законом розподілу

X	-3	-2	2	3
p	0,3	0,1	0,2	0,4

Знайти: 1) Закон розподілу $Y = X^2$;

2) математичне сподівання $M(X)$;

3) математичне сподівання $M(Y)$.

Розв'язування. 1) Знайдемо можливі значення Y:

$$y_1 = {x_1}^2 = (-3)^2 = 9;\ \ y_2 = {x_2}^2 = (-2)^2 = 4;\ \ y_3 = {x_3}^2 = 2^2 = 4;\ \ y_4 = {x_4}^2 = 3^2 = 9.$$

Знайдемо ймовірності можливих значень Y. Для того, щоб дискретна випадкова величина Y набула значення $Y = 4$, достатньо , щоб випадкова величина X набула значення $X = -2$ або $X = 2$. Останні дві події несумісні, а їх ймовірності відповідно дорівнюють 0,1 і 0,2. Тому ймовірність події $Y = 4$ за теоремою додавання рівна:

$$P(Y = 4) = P(X = -2) + P(X = 2) = 0{,}1 + 0{,}2 = 0{,}3.$$

Аналогічно знайдемо ймовірність можливого значення $Y = 9$:

$$P(Y = 9) = P(X = -3) + P(X = 3) = 0{,}3 + 0{,}4 = 0{,}7.$$

Запишемо шуканий закон розподілу:

X	4	9
p	0,3	0,7

2) $M(X) = -3\cdot 0{,}3 - 2\cdot 0{,}1 + 2\cdot 0{,}2 + 3\cdot 0{,}4 = 2\cdot 0{,}1 + 3\cdot 0{,}1 = 5\cdot 0{,}1 = 0{,}5.$

3) $M(Y) = 4\cdot 0{,}3 + 9\cdot 0{,}7 = 7{,}4.$

Приклад 57. Дискретна випадкова величина X набуває трьох можливих значень: $x_1 = 4$ з імовірністю $p_1 = 0{,}5;$ $x_2 = 6$ з імовірністю $p_2 = 0{,}3;$ x_3 з імовірністю p_3. Знайти x_3 та p_3 , знаючи , що $M(X) = 8.$

Розв'язування. Для знаходження двох невідомих x_3 та p_3 маємо два рівняння :

$$\begin{cases} p_1 + p_2 + p_3 = 1 \\ x_1 p_1 + x_2 p_2 + x_3 p_3 = M(X) \end{cases},$$

або

$$\begin{cases} p_3 = 1 - p_1 - p_2 \\ x_3 = \dfrac{1}{p_3}(M(X) - x_1 p_1 - x_2 p_2). \end{cases}$$

Після підстановки в ці рівняння даних умови задачі дістанемо:

$$p_3 = 1 - 0{,}5 - 0{,}3 = 0{,}2; x_3 = \frac{1}{0{,}2}(8 - 2 - 1{,}8) = \frac{4{,}2}{0{,}2} = 21.$$

Приклад 58. (*Для самостійної опрацювання*). Незалежні дискретні випадкові величини X та Y задано законами розподілів:

X	0	2	3
P	0,5	0,3	0,2

Y	1	2	3
P	0,2	0,4	0,4

Скласти закон розподілу величин 1) $Z = X + Y$; 2) $V = X \cdot Y$.

§6. Дисперсія дискретної випадкової величини та її властивості

Означення. ***Дисперсією*** $D(X)$ ***(розсіюванням) випадкової величини*** X ***називається число, яке дорівнює математичному сподіванню квадрата відхилення випадкової величини*** X ***від її математичного сподівання:***

$$D(X) = M[(X - M(X))^2]. \tag{44}$$

Приклад 59. Знайти дисперсію випадкової величини $X -$ кількості появи гербів при підкиданні трьох монет.

Розв'язування. Як відомо з прикладу 51, закон розподілу X має вид:

X	*0*	*1*	*2*	*3*
p	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

і $M(X)$ = 1,5. Запишемо розподіли випадкових величин $X - M(X)$ і $(X - M(X))^2$:

$X - M(X)$	*-1,5*	*-0,5*	*0,5*	*1,5*
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$(X - M(X))^2$	*0,25*	*2,25*
P	$\frac{6}{8}$	$\frac{2}{8}$

Отже, згідно з рівністю (44), знаходимо:

$$D(X) = 2{,}25 \cdot \frac{2}{8} + 0{,}25 \cdot \frac{6}{8} = \frac{1}{8} \cdot 6 = 0{,}75.$$

Наведемо важливу формулу для обчислення дисперсії випадкової величини.

Дисперсія дискретної випадкової величини X ***дорівнює різниці між математичним сподіванням квадрата випадкової величини*** X ***та квадратом її математичного сподівання:***

$$D(X) = M(X^2) - (M(X))^2. \tag{45}$$

Дійсно, згідно з властивостями $M(X)$, знаходимо:

$$D(X) = M[(X - M(X))^2] = M[X^2 - 2MX(X) + M^2(X)] =$$
$$= M(X^2) - 2M(X)M(X) + M^2(X) = M(X^2) - M^2(X).$$

Приклад 60. Знайти дисперсію випадкової величини за формулою (45) у прикладі 59.

Розв'язування. Запишемо розподіл випадкової величини X^2. Зауважимо, що всі значення X^2 знаходимо шляхом піднесення до квадрату відповідних значень X, а ймовірності цих значень не змінюються.

X^2	*0*	*1*	*4*	*9*
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Тоді,

$$M(X^2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 3\ ;\ M^2(X) = (1{,}5)^2 = 2{,}25.$$

Отже,

$$D(X) = 3 - 2{,}25 = 0{,}75.$$

Отримали той самий результат, що і за формулою (44).

Розглянемо властивості дисперсії дискретної випадкової величини X.

1. ***Дисперсія*** X ***невід'ємна, тобто*** $D(X) \geq 0$.

Дійсно, величина $(X - M(X))^2$ невід'ємна, тому згідно з рівністю (44), враховуючи, що $p_k \geq 0 (k = \overline{1,n})$, $D(X) \geq 0$.

2. ***Дисперсія сталої величини*** C ***дорівнює нулю, тобто*** $D(C) = 0$.

Дійсно, $M(C) = C$, тому $C - M(C) = 0$, отже, $D(C - C) = 0$.

3. Сталий множник можна виносити за знак дисперсії, піднісши його до квадрату : $D(CX) = C^2 D(X)$.

Дійсно,

$$D(CX) = M(C^2 X^2) - M^2(CX) = C^2 M(X^2) - (CM(X))^2 =$$
$$= C^2 (M(X^2) - M^2(X)) = C^2 D(X).$$

4. Дисперсія алгебраїчної суми двох незалежних дискретних випадкових величин X ***та*** Y ***дорівнює сумі їх дисперсії, тобто***

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

Доведення. Доведемо спочатку для суми $X + Y$. Згідно з рівністю (44) і властивостями $M(X)$ маємо :

$$D(X + Y) = M(X + Y)^2 - (M(X + Y))^2 = M(X^2 + 2XY + Y^2) - (M(X) + M(Y))^2 =$$
$$= M(X^2) + 2M(X)M(Y) + M(Y^2) - M^2(X) - 2M(X)M(Y) - M^2(Y) =$$
$$= M(X^2) - M^2(X) + M(Y^2) - M^2(Y) = D(X) - D(Y).$$

Аналогічно: $D(X - Y) = D(X + (-1)\cdot Y) = D(X) + D((-1)\cdot Y) = D(X) + D(Y)$.

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

$$D(X + Y + Z) = D(X + Y) + D(Z) = D(X) + D(Y) + D(Z).$$

Наслідок 2. Дисперсія суми сталої величини C і випадкової величини X дорівнює дисперсії випадкової величини X.

Дійсно, $D(X + C) = D(X) + D(C) = D(X) + 0 = D(X)$.

Зауваження. Наслідок 2 означає , що зміщення випадкової величини на сталу не змінює її дисперсії.

Приклад 61. За даними можливих прибутків (у тис. грн.) для двох видів інвестування та їх ймовірностей:

	Порівняння варіантів						
чистий прибуток, тис.грн.	*-3*	*-2*	*-1*	*0*	*1*	*2*	*3*
ймовірність інвестиції 1	*0*	*0*	*0,1*	*0,2*	*0,3*	*0,2*	*0,2*
ймовірність інвестиції 2	*0,1*	*0,1*	*0,1*	*0*	*0,2*	*0,1*	*0,2*

визначити, яка з інвестицій краща і чому.

Розв'язування. Знайдемо очікуваний прибуток для кожного варіанту інвестицій :

$$M(X_1) = (-3\times 0) + (-2\times 0) + (-1\times 0{,}1) + (0\times 0{,}2) + (1\times 0{,}3) + (2\times 0{,}2) + (3\times 0{,}2) + (4\times 0) =$$
$$= -0{,}1 + 0{,}3 + 0{,}4 + 0{,}6 = 1{,}2 \text{ (тис. грн.)}.$$
$$M(X_2) = (-3\times 0{,}1) + (-2\times 0{,}1) + (-1\times 0{,}1) + (0\times 0) + (1\times 0{,}2) + (2\times 0{,}1) + (3\times 0{,}2) + (4\times 0{,}2) =$$
$$= -0{,}3 - 0{,}2 - 0{,}1 + 0{,}2 + 0{,}2 + 0{,}6 + 0{,}8 = 1{,}2 \text{ (тис. грн.)}.$$

Якщо брати до уваги лише очікуваний прибуток, то інвестиції рівноцінні. З'ясуємо який ризик пов'язаний з даними інвестиціями. Для цього підрахуємо відповідні дисперсії та середні квадратичні відхилення:

$$D(X_1) = 1{,}56 \text{ тис. грн.}, \ \delta(X_1) = \sqrt{1{,}56} = 1{,}25 \text{ тис. грн.},$$
$$D(X_2) = 5{,}56 \text{ тис. грн.}, \ \delta(X_2) = \sqrt{5{,}56} = 2{,}36 \text{ тис. грн..}$$

Таким чином, ризикованішою є інвестиція 2, оскільки її дисперсія (розсіювання) є більшим. Тобто ця інвестиція є менше прогнозованою.

Обчислення дисперсії дискретних випадкових величин
Біномний розподіл

Дисперсія біномного розподілу дискретної випадкової величини X з параметрами n і p дорівнює npq, тобто $D(X) = npq.$ Дійсно:

$$M(X^2) = \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} = np\sum_{k=0}^{n} \frac{k^2}{n} C_n^k p^k q^{n-k} = np\sum_{k=0}^{n} kC_{n-1}^{k-1} p^{k-1} q^{n-k} =$$
$$= np(\sum_{k=0}^{n} (k-1)C_{n-1}^{k-1} p^{k-1} q^{n-k} + \sum_{k=0}^{n} C_{n-1}^{k-1} p^{k-1} q^{n-1}) =$$
$$= np((n-1)p+1) = n^2p^2 - np^2 + np.$$

Таким чином, згідно з рівністю (45), знаходимо:

$$D(X) = M(X^2) - M^2(X) = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) = npq.$$

Розподіл Пуассона.

Покажемо, що дисперсія розподілу Пуассона с параметром λ дорівнює λ, тобто: $D(X)=\lambda$.

Дійсно, за умовою випадкова величина X має розподіл Пуассона, тому $M(X)=\lambda$. Для цього спочатку знаходимо:

$$M(X(X-1))=\sum_{л=0}^{\infty}k(k-1)\frac{\lambda^k}{k!}e^{-\lambda}=\lambda^2\sum_{k=0}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda}=\lambda^2e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}=\lambda^2e^{-\lambda}e^{\lambda}=\lambda^2.$$

Враховуючи властивості математичного сподівання, маємо

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

Підставимо одержані значення в останню рівність і знайдемо $M(X^2)$.

$$\lambda^2=M(X^2)-\lambda, \text{звідки } M(X^2)=\lambda^2+\lambda..$$

Отже,

$$D(X)=M(X^2)-M^2(X)=\lambda^2+\lambda-\lambda^2=\lambda.$$

Геометричний розподіл

Доведемо, що дисперсія геометричного розподілу дискретної випадкової величини X дорівнює $\frac{q}{p^2}$, де p $(0<p<1)$ і q $(0<q<1)$ – ймовірності відповідно появи і не появи події A у кожному випробуванні.

Дійсно. За умовою розподіл геометричний, тому $M(X)=\frac{1}{p}.$

Знайдемо $M(X^2)$:

$$\begin{aligned}M(X^2)&=1^2\cdot p+2^2\cdot qp+3^2\cdot q^2p+\cdots+n^2q^{n-1}p+\cdots=\\&=p(1^2+2^2q+3^2q^2+\cdots+n^2q^{n-1}+\cdots)=\\&=p\frac{1+q}{(1-q)^3}=\frac{p(1+q)}{p^3}=\frac{1+q}{p^2}.\end{aligned}$$

Далі згідно з рівністю (45) знаходимо:

$$D(X)=M(X^2)-M^2(X)=\frac{1+q}{p^2}-\frac{1}{p^2}=\frac{q}{p}.$$

Середнє квадратичне відхилення дискретної випадкової величини

Для оцінки розсіювання можливих значень випадкової величини X навколо її середнього значення $M(X)$ вводять числову характеристику, яка має ту саму розмірність, що і випадкова величина (на відміну від дисперсії, яка вимірюється у квадратних щодо X одиницях). Це приводить до введення середнього квадратичного відхилення, зміст якого такий же, як у дисперсії, а розмірність – як у випадкової величини.

Означення. ***Середнім квадратичним відхиленням дискретної випадкової величини*** X ***називається число*** $\delta(X)$ ***, що визначається рівністю:***

$$\delta(X)=\sqrt{D(X)}. \tag{46}$$

Приклад 62. Знайти середнє квадратичне відхилення у прикладі 60.

Розв'язування. Згідно з формулою (46) маємо:

$$\delta(X) = \sqrt{D(X)} = \sqrt{3}\big/2 \approx 0{,}865.$$

Приклад 63. Задано закон розподілу дискретної випадкової величини:

X	0	1	2	3
p	$0{,}504$	$0{,}398$	$0{,}092$	$0{,}006$

Знайти математичне сподівання та середнє квадратичне відхилення.

Розв'язування. За формулами (33) – (35) знаходимо:

$M(X) = 0 \cdot 0{,}504 + 1 \cdot 0{,}398 + 2 \cdot 0{,}092 + 3 \cdot 0{,}006 = 0{,}6;$

$M(X^2) = 0^2 \cdot 0{,}504 + 1^2 \cdot 0{,}398 + 2^2 \cdot 0{,}092 + 3^2 \cdot 0{,}006 = 0{,}82;$

$D(X) = M(X^2) - (MX)^2 = 0{,}82 - (0{,}6)^2 = 0{,}46;\quad \sigma(X) = \sqrt{D(X)} = \sqrt{0{,}46} \approx 0{,}678.$

Має місце важлива властивість середнього квадратичного відхилення.

Середнє квадратичне відхилення суми скінченого числа взаємно незалежних випадкових величин дорівнює квадратному кореню з суми квадратів їх середніх квадратичних відхилень, тобто:

$$\delta\left(\sum_{k=1}^{n} X_k\right) = \sqrt{\sum_{k=1}^{n} \delta^2(X_k)}\,. \tag{47}$$

Дійсно, якщо $X = \sum_{k=1}^{n} X_k$ і X_i – взаємно незалежні, то $D(X) = \sum_{k=1}^{n} D(X_k)$.

Тоді:

$$\delta\left(\sum_{k=1}^{n} X_k\right) = \sqrt{D(X)} = \sqrt{\sum_{k=1}^{n} D(X_k)} = \sqrt{\delta^2(X_k)}.$$

Незалежні випадкові величини, що мають однаковий розподіл

Випадкові величини , що мають однаковий розподіл, мають і однакові числові характеристики. Нехай задані взаємно незалежні випадкові величини $X_1, X_2, \ldots, X_n$, що мають однаковий розподіл (позначимо їх середнє арифметичне через $\overline{X}$, тобто $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$), і для яких $M(X_k) = a, D(X_k) = D, \delta(X_k) = \delta \quad (k = \overline{1,n})$.

Тоді мають місце наступні властивості:

1. ***Математичне сподівання середнього арифметичного однаково розподілених взаємно незалежних випадкових величин дорівнює математичному сподіванню кожної з них:*** $M(\overline{X}) = a.$

   Дійсно:

$$M(\overline{X}) = M\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} M(X_k) = a.$$

2. ***Дисперсія середнього арифметичного*** n ***однаково розподілених взаємно незалежних випадкових величин у*** n ***разів менша за дисперсію кожної з них:***

$$D(\overline{X}) = \frac{D(X)}{n} = \frac{D}{n}.$$

Дійсно:

$$D(\overline{X}) = D\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} D(X_k) = \frac{D \cdot n}{n^2} = \frac{D}{n}.$$

3. ***Середнє квадратичне відхилення середнього арифметичного*** n ***однаково розподілених взаємно незалежних випадкових величин у*** $\sqrt{n}$ ***разів менша за середнє квадратичне відхилення кожної з них:*** $\delta(\overline{X}) = \frac{\delta}{\sqrt{n}}$.

Дійсно:

$$\delta(\overline{X}) = \sqrt{D(\overline{X})} = \sqrt{\frac{D}{n}} = \frac{\sqrt{D}}{\sqrt{n}} = \frac{\delta}{\sqrt{n}}.$$

Зауваження. Аналізуючи наведені властивості, робимо висновок, що середнє арифметичне досить великої кількості взаємно незалежних випадкових величин має значно менше розсіювання, ніж кожна окрема випадкова величина.

Моменти розподілу дискретної випадкової величини

Означення. ***Початковим моментом*** k***-го порядку дискретної випадкової величини*** X ***називається число*:**

$$v_k = M(X^k)(k = \overline{1,n}). \tag{48}$$

Очевидно, що $v_1 = M(X), v_2 = M(X^2),\ D(X) = M(X^2) - M^2(X) = v_2 - v_1^2$.

Означення. ***Центральним моментом*** k***-го порядку дискретної випадкової величини*** X ***називається число:***

$$\mu_k = M(X - M(X))^k (k = \overline{1,n}). \tag{49}$$

Очевидно, що $\mu_2 = M(X - M(X))^2 = D(X) = v_2 - v_1^2$.

Зауваження 1. Початкові та центральні моменти при $k \geq 2$ дозволяють краще враховувати вплив на математичне сподівання досить великих можливих значень випадкової величини X, які мають малу ймовірність.

Приклад 64. Знайти початкові та центральні моменти першого та другого порядків дискретної випадкової величини X, заданої розподілом :

X	*1*	*2*	*5*	*100*
P	*0,6*	*0,2*	*0,19*	*0,01*

Розв'язування. Згідно з рівністю (48) знаходимо:

$$v_1 = M(X) = 1 \cdot 0{,}6 + 2 \cdot 0{,}2 + 5 \cdot 0{,}19 + 100 \cdot 0{,}01 = 2{,}95.$$

Запишемо закон розподілу X^2:

X^2	*1*	*4*	*25*	*10000*
P	*0,6*	*0,2*	*0,19*	*0,01*

Далі обчислимо v_2 і μ_2:

$$v_2 = M(X^2) = 1 \cdot 0{,}6 + 4 \cdot 0{,}2 + 25 \cdot 0{,}19 + 10000 \cdot 0{,}01 = 106{,}15,$$

$$\mu_2 = \nu_2 - \nu_1^{\,2} = 106{,}15 - 2{,}95 = 103{,}2.$$

Зауважимо, що величина $M(X^2)$ значно більша за $M(X)$.

Приклад 65. Знайти дисперсію та середнє квадратичне відхилення випадкової величини X, заданої законом розподілу :

X	0	1	2	3	4	5
P	$0{,}40$	$0{,}25$	$0{,}02$	$0{,}10$	$0{,}01$	$0{,}22$

Розв'язування. Для обчислення дисперсії скористаємось формулою $D(X) = M(X^2) - (M(X))^2.$

Послідовно обчислюємо:

1) $M(X) = 4\cdot 0{,}01 + 2\cdot 0{,}02 + 1\cdot 0{,}25 + 0\cdot 0{,}40 + 3\cdot 0{,}10 + 5\cdot 0{,}22 = 1{,}73;$

2) $M^2(X) = 1{,}73^2 = 2{,}99;$

3) $M(X^2) = 16\cdot 0{,}01 + 4\cdot 0{,}02 + 1\cdot 0{,}25 + 0\cdot 0{,}40 + 9\cdot 0{,}10 + 25\cdot 0{,}22 = 6{,}89;$

4) $D(X) = 6{,}89 - 2{,}99 = 3{,}90;$

5) $\delta(X) = \sqrt{D(X)} = \sqrt{3{,}90} = 1{,}97.$

Приклад 66. Дискретні величини X та Y незалежні і мають дисперсії $D(X) = 5, D(Y) = 10.$ Обчислити дисперсію та середнє квадратичне відхилення випадкової величини $Z = 2X + 3Y + 7.$

Розв'язування. Оскільки величини X та Y незалежні, то незалежні також і величини $2X$ та $3Y$. Використовуючи властивості дисперсії, одержимо:

$$D(Z) = D(2X + 3Y + 7) = D(2X) + D(3Y) + D(7) =$$
$$= 2^2 D(X) + 3^2 D(Y) + D(7) = 4\cdot 5 + 9\cdot 10 + 0 = 110;$$
$$\delta(Z) = \sqrt{110} \approx 10{,}488.$$

Приклад 67. Проводиться 1000 незалежних повторних випробувань, у кожному з яких ймовірність появи деякої події дорівнює 0,6. Визначити математичне сподівання та дисперсію числа появи події в цих випробуваннях.

Розв'язування. Дискретна величина X – число появи деякої події в 1000 незалежних випробуваннях – розподілена за біномним законом. За умовою $n = 1000, p = 0{,}6, q = 0{,}4.$

$$M(X) = n\cdot p = 1000\cdot 0{,}6 = 600; \quad D(X) = npq = 1000\cdot 0{,}6\cdot 0{,}4 = 240.$$

Завдання для самостійної роботи студентів

1. Ймовірність виграти по лотерейному білету дорівнює 0,4. Придбано 30 білетів. Визначити математичне сподівання та дисперсію числа білетів, на які випадають виграші.
2. Ймовірність порушення герметичності банки в партії консервів дорівнює 0,02. Визначити математичне сподівання та дисперсію числа банок з порушенням герметичності у партії із 1000 банок консервів.

3. Комплекс по виробництву вершкового масла складається з двох незалежно працюючих ліній. Ймовірності відмови кожної з ліній протягом місяця однакові. Знайти дисперсію дискретної випадкової величини X - числа ліній, що відмовили протягом місяця, якщо $M(X) = 0,20$.
4. Необхідно дослідити 1000 проб сировини. Ймовірність того, що сировина виявиться неякісною, в кожній пробі одна й та ж сама. Визначити дисперсію дискретної випадкової величини X - числа проб, в яких сировина виявиться неякісною, якщо $M(X) = 200$.
5. Проводяться незалежні випробування з однаковою ймовірністю появи події A в кожному випробуванні. Знайти ймовірність появи події A, якщо дисперсія числа появи події в трьох незалежних випробуваннях дорівнює 0,63.
6. Ймовірність того, що витрата води на підприємстві не перевищить норми в робочий день, p - стала величина. Дисперсія дискретної випадкової величини X - числа днів, протягом яких відбудеться перевитрата води з перших десяти днів місяця, - дорівнює 2,4. Знайти p.
7. Пекарня складається з n агрегатів . Ймовірність відмови кожного агрегату в одному випробуванні дорівнює p. Знайти математичне сподівання та дисперсію числа таких випробувань, у кожному з яких відмовить рівно k агрегатів, якщо всього проведено N випробувань. Припускається, випробування незалежні одне від одного.
8. Відділ технічного контролю перевіряє вироби на стандартність. Ймовірність того, що виріб стандартний, дорівнює 0,9. У кожній партії міститься 5 виробів. Знайти математичне сподівання і дисперсію дискретної випадкової величини X - числа партій, у кожній з яких виявиться рівно 4 стандартні вироби, якщо перевірці підлягають 50 партій.
9. Проводяться багаторазові випробування технологічної лінії на надійність доти, доки лінія не відмовить. Знайти математичне сподівання та дисперсію дискретної випадкової величини X - числа дослідів, які треба провести . Ймовірність відмови лінії в кожному досліді дорівнює 0,1.
10. Ймовірність того, що стрілець влучить у мішень при одному пострілі, дорівнює 0,8. Стрільцю видаються патрони до того часу, доки він не промахнеться. Знайти математичне сподівання та дисперсію дискретної випадкової величини X - числа патронів, виданих стрільцю.

§ 7. *Числові характеристики неперервних випадкових величин*

Вище були вивчені числові характеристики для дискретних випадкових величин. Перейдемо до розгляду числових характеристик неперервних випадкових величин.

Математичне сподівання неперервної випадкової величини

Нехай неперервна випадкова величина *X* задана щільністю розподілу *f(x)*.

Означення. ***Якщо неперервна випадкова величина X має щільність f(x) , то її математичне сподівання знаходиться за формулою***

$$M(X) = \int_{-\infty}^{+\infty} xf(x)dx, \qquad (50)$$

за умови, що невласний інтеграл у (50) є абсолютно збіжним.

Зауваження 1. *Якщо можливі значення X належать відрізку [a,b], то її математичне сподівання знаходиться за формулою:*

$$M(X) = \int_{a}^{b} xf(x)dx. \qquad (51)$$

Дисперсія неперервної випадкової величини

Означення. Дисперсія неперервної випадкової величини визначається, як і для дискретної випадкової величини, формулою:

$$D(X) = M[(X - M(X))^2]. \qquad (52)$$

Зауваження 2. ***Якщо неперервна випадкова величина X набуває значень з відрізку [a,b] , то дисперсія*** обчислюється згідно співвідношення:

$$D(X) = \int_{a}^{b} (x - M(X))^2 f(x)dx. \qquad (53)$$

Зауваження 3. ***Якщо можливі значення випадкової величини X належать R, то***

$$D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 f(x)dx. \qquad (54)$$

Зауваження 4. На практиці дисперсію неперервної випадкової величини *X* обчислюють за формулами:

$$D(X) = \int_{a}^{b} x^2 f(x)dx - M^2(X), \qquad (55)$$

або відповідно,

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - M^2(X). \qquad (56)$$

Середнє квадратичне відхилення

Означення. Середнє квадратичне відхилення неперервної випадкової величини визначається згідно формули:

$$\sigma(X) = \sqrt{D(X)}. \qquad (57)$$

Зауваження 5. Числові характеристики неперервних випадкових величин мають властивості, аналогічні властивостям числових характеристик дискретних випадкових величин.

Приклад 68. Знайти числові характеристики неперервної випадкової величини *X* з функцією розподілу:

$$F(x)=\begin{cases}0 \quad при \ x\le 0\\ \dfrac{x^2}{4} \ \ при \ 0<x\le 2\\ 1 \quad при \ x>2\end{cases}.$$

Розв'язування. Знайдемо щільність розподілу ймовірностей:

$$f(x)=F'(x)=\begin{cases}0 \quad при \ x\le 0\\ \dfrac{x}{2} \ \ при \ 0<x\le 2\\ 0 \quad при \ x>2\end{cases}.$$

Математичне сподівання знайдемо за формулою (51):

$$M(X)=\int_0^2 x\frac{x}{2}dx=\frac{x^3}{6}\Big|_0^2=\frac{4}{3} \ .$$

Дисперсію знайдемо за (55):

$$D(X)=\int_0^2 x^2\frac{x}{2}dx-(\frac{4}{3})^2=\frac{x^4}{8}\Big|_0^2-\frac{16}{9}=\frac{2}{9}.$$

Середнє квадратичне відхилення обчислюємо за формулою (57):

$$\sigma(X)=\sqrt{\frac{2}{9}}=\frac{\sqrt{2}}{3}\approx 0{,}4714 .$$

Приклад 69. Задана неперервна випадкова величина з щільністю розподілу ймовірностей:

$$f(x)=\begin{cases}A\cos x, \text{при} -{\pi}/{2}\le x\le {\pi}/{2}\\ 0, \text{при} |x|>{\pi}/{2}\end{cases}.$$

Знайти: а) величину сталої *A*; б) функцію розподілу *F*(*x*); в) математичне сподівання та дисперсію випадкової величини; г) ймовірність того, що випадкова величина X набуде значеннь з проміжку $0\le X<{\pi}/{2}$.

Розв’язування. Для знаходження значення *A* використаємо властивість щільності:

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

Тому,

$$1=\int_{-\infty}^{+\infty} f(x)dx=\int_{-\pi/2}^{\pi/2} A\cos x dx=A\sin x\Big|_{-\pi/2}^{\pi/2}=A\left(\sin\frac{\pi}{2}-\sin\left(-\frac{\pi}{2}\right)\right)=2A; \ A={1}/{2} .$$

Отже, щільність розподілу ймовірностей набуде вигляду:

$$f(x)=\begin{cases}\frac{1}{2}\cos x, \text{при } -\frac{\pi}{2}\le x\le\frac{\pi}{2}\\ 0, \text{при } |x|>\frac{\pi}{2}\end{cases}.$$

Оскільки $F(x)=\int\limits_{-\infty}^{x} f(t)dt$, то знайдемо значення функції розподілу на кожному інтервалі окремо:

1. $x\le -\frac{\pi}{2}$, $\quad F(x)=\int\limits_{-\infty}^{x}0dt=0$;

2. $-\frac{\pi}{2}<x\le\frac{\pi}{2}$,

$$F(x)=\int\limits_{-\infty}^{-\pi/2}0dt+\int\limits_{-\pi/2}^{x}\frac{1}{2}\cos tdt=\frac{1}{2}\sin t\Big|_{-\pi/2}^{x}=\frac{1}{2}\left(\sin x-\sin(-\frac{\pi}{2})\right)=\frac{1}{2}(\sin x+1);$$

3. $x>\frac{\pi}{2}$, $F(x)=\int\limits_{-\infty}^{-\pi/2}0dt+\int\limits_{-\pi/2}^{\pi/2}\frac{1}{2}\cos tdt+\int\limits_{\pi/2}^{x}0dt=\frac{1}{2}\sin t\Big|_{-\pi/2}^{\pi/2}=\frac{1}{2}\left(\sin\frac{\pi}{2}-\sin\left(-\frac{\pi}{2}\right)\right)=1.$

Отже, шукана функція розподілу для заданої неперервної випадкової величини є такою:

$$F(x)=\begin{cases}0, & x\le-\frac{\pi}{2}\\ \frac{1}{2}(\sin x+1), & -\frac{\pi}{2}<x\le\frac{\pi}{2}\\ 1, & x>\frac{\pi}{2}\end{cases}.$$

Знаходимо математичне сподівання та дисперсію:

$$M(X)=\int\limits_{-\infty}^{+\infty}xf(x)dx=\frac{1}{2}\int\limits_{-\pi/2}^{\pi/2}x\cos xdx=\begin{vmatrix}u=x & du=dx\\ dv=\cos xdx & v=\sin x\end{vmatrix}=\frac{1}{2}\left(x\sin x\Big|_{-\pi/2}^{\pi/2}-\int\limits_{-\pi/2}^{\pi/2}\sin xdx\right)=$$

$$=\frac{1}{2}\left(\frac{\pi}{2}\sin\left(\frac{\pi}{2}\right)-\left(-\frac{\pi}{2}\right)\sin\left(-\frac{\pi}{2}\right)+\cos x\Big|_{-\pi/2}^{\pi/2}\right)=0;$$

$$M(X^2)=\int\limits_{-\infty}^{+\infty}x^2f(x)dx=\int\limits_{-\pi/2}^{\pi/2}x^2\frac{1}{2}\cos xdx=\begin{vmatrix}u=x^2 & du=2xdx\\ dv=\frac{1}{2}\cos xdx & v=\frac{1}{2}\sin x\end{vmatrix}=\frac{1}{2}x^2\sin x\Big|_{-\pi/2}^{\pi/2}-\int\limits_{-\pi/2}^{\pi/2}x\sin xdx=$$

$$=\begin{vmatrix}u=x & du=dx\\ dv=\sin xdx & v=-\cos x\end{vmatrix}=\frac{\pi^2}{8}\sin\left(\frac{\pi}{2}\right)-\frac{\pi^2}{8}\sin\left(-\frac{\pi}{2}\right)-\left(-x\cos x\Big|_{-\pi/2}^{\pi/2}-\int\limits_{-\pi/2}^{\pi/2}(-\cos x)dx\right)=$$

$$= \pi^2/4 - \sin x \Big|_{-\pi/2}^{\pi/2} = \pi^2/4 - \left(\sin \pi/2 - \sin\left(-\pi/2\right)\right) = \pi^2/4 - 2;$$

$$D(X) = M\left(X^2\right) - (MX)^2 = \pi^2/4 - 2.$$

Використовуючи функцію розподілу, знаходимо шукану ймовірність:

$$P\left\{0 \le X < \pi/6\right\} = F\left(\pi/6\right) - F(0) = \frac{1}{2}\left(\sin\left(\pi/6\right) + 1\right) - \frac{1}{2}\left(\sin 0 + 1\right) = \frac{1}{4}.$$

Завдання для самостійної роботи студентів

1. Випадкова величина задана функцією розподілу ймовірностей

$$F(x) = \begin{cases} 0, & x \le 0 \\ \sin x, & 0 < x \le \dfrac{\pi}{2} \\ 1, & x > \dfrac{\pi}{2} \end{cases}.$$

Знайти щільність розподілу *f(x)* та побудувати графіки функцій $F(x)$, $f(x)$.

2. Випадкова величина задана функцією розподілу

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

Знайти щільність розподілу *f(x)* та побудувати графіки функцій $F(x)$, $f(x)$.

3. Неперервна випадкова величина X задана щільністю розподілу ймовірностей:

$$f(x) = \begin{cases} 0, & x \le \dfrac{\pi}{3} \\ 3\sin 3x, & \dfrac{\pi}{6} < x \le \dfrac{\pi}{3} \\ 0, & x > \dfrac{\pi}{3} \end{cases}$$

Знайти ймовірність того, що випадкова величина X набуде значення з інтервалу $(-\pi, \pi)$.

4. Неперервна випадкова величина X задана щільністю розподілу ймовірностей:

$$f(x) = \begin{cases} 0, & x \le 1 \\ \ln x, & 1 < x \le e \\ 0, & x > e \end{cases}.$$

Визначити ймовірність того, що випадкова величина X набуде значення з інтервалу *(1; e)*.

5. Задано щільність розподілу ймовірностей випадкової величини *X:*

$$f(x)=\begin{cases}0, & x\le 1\\ Ax-\dfrac{1}{2}, & 1<x\le 2\\ 0, & x>2\end{cases}.$$

Знайти: 1) сталий параметр *A;* 2) функцію розподілу випадкової величини *X;* 3) побудувати графіки функцій $F(x)$, $f(x)$.

6. Для заданої щільності розподілу ймовірностей $f(x)=\dfrac{a}{1+x^2}$ при $-\infty<x<+\infty$, визначити: 1) сталий параметр *a;* 2) ймовірність попадання неперервної випадкової величини в інтервал (-2,4); 3) функцію розподілу ймовірностей випадкової величини *X.*

7. Неперервна випадкова величина *X* задана щільністю розподілу ймовірностей:

$$f(x)=\begin{cases}0, & x\le -1\\ \dfrac{A}{\sqrt{1-x^2}}, & -1<x\le 1\\ 0, & x>1\end{cases}$$

Знайти: 1) сталий параметр *A*; 2) функцію розподілу випадкової величини *X;* 3) побудувати графіки функцій $F(x)$, $f(x)$.

8. Випадкова величина *X* задана функцією розподілу

$$F(x)=\begin{cases}0, & x\le 1\\ \dfrac{x^2-x}{2}, & 1<x\le 2\\ 1, & x>2\end{cases}.$$

Обчислити а) щільність розподілу випадкової величини *X*;

б) математичне сподівання та дисперсію, середнє квадратичне відхилення;

в) $P\{0{,}5\le X<1{,}5\}$, $P\{1{,}5\le X<2{,}5\}$.

9. Випадкова величина *X* задана функцією розподілу

$$F(x)=\begin{cases}0, & x\le 0\\ x^3, & 0<x\le 1\\ 1, & x>1\end{cases}.$$

Знайдіть: а) щільність розподілу випадкової величини *X*;

б) математичне сподівання та дисперсію, середнє квадратичне відхилення;

в) $P\{0\le X<0{,}5\}$, $P\{0{,}3\le X<0{,}6\}$.

10. Задана щільність розподілу випадкової величини *X*:

$$f(x)=\begin{cases}0, & x\le 2\\ a(x-2), & 2<x\le 3\\ 0, & x>3\end{cases}.$$

Знайдіть: а) коефіцієнт a;

б) математичне сподівання, дисперсію, середнє квадратичне відхилення;

в) функцію розподілу випадкової величини X.

11. Випадкова величина X задана функцією розподілу

$$F(x)=\begin{cases}0, & x\le 0\\ \frac{2}{\pi}arctgx, & x>0\end{cases}$$

Знайти ймовірність того, що величина X набуде значення з інтервалу (1; 3).

12. Випадкова величина X задана функцією розподілу

$$F(x)=\begin{cases}0, & x\le 0\\ x(2-x), & 0<x\le 1\\ 1, & x>1\end{cases}.$$

Знайти ймовірність того, що неперервна випадкова величина X набуде значення з інтервалу $(0.5, 1)$.

13. Випадкова величина X задана функцією розподілу

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

Знайти ймовірність того, що неперервна випадкова величина X прийме значення: 1) менше 0,5; 2) менше 1,5; 3) не менше 2,5; 4) не менше 4.

14. Випадкова величина X задана функцією розподілу

$$F(x)=\begin{cases}0, & x\le -1\\ \frac{1}{2}+\frac{1}{\pi}arcsin\,x, & -1<x\le 1\\ 1, & x>1\end{cases}.$$

Знайти ймовірність того, що випадкова величина X прийме значення з інтервалу *(0.5;0.8)*. Побудувати графік функції розподілу $F(x)$.

15. Неперервна випадкова величина X задана щільністю розподілу $f(x)=3x^2$ на інтервалі (0;1), а зовні його - $f(x)=0$. Знайти числові характеристики неперервної випадкової величини X.

16. Рівномірно розподілена випадкова величина X задана щільністю розподілу $f(x)=\frac{1}{5}$ на інтервалі (1;6), а зовні його - $f(x)=0$. Знайти числові характеристики неперервної випадкової величини X.

17. Випадкова величина X задана на всій вісі Ox функцією розподілу :

$$F(x)=\frac{1}{2}+\frac{1}{\pi}arctg\ x.$$

За відомою ймовірністю Р=$\frac{1}{3}$ знайти значення x_0 таке, що випадкова величина X набуде значення більше ніж x_0.

§ 8. Види законів розподілу неперервних випадкових величин

При розв'язуванні багатьох прикладних задач статистичного аналізу важливими є закони розподілу неперервних випадкових величин, зокрема рівномірний, показниковий, нормальний. Зупинимося більш детально на них.

Рівномірний закон розподілу

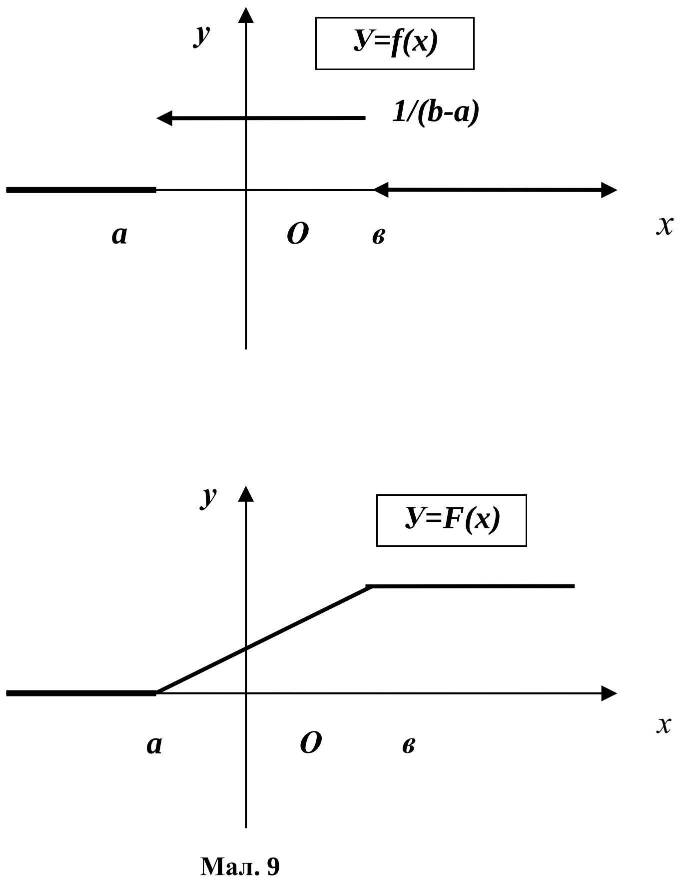
Означення. ***Неперервна випадкова величина X називається рівномірно розподіленою на відрізку*** $[a,b]$***, якщо щільність розподілу її має вигляд:***

$$f(x)=\begin{cases}0, & x\le a\\ \dfrac{1}{b-a}, & a<x\le b\\ 0, & x>b\end{cases}. \qquad (58)$$

Відповідна функція розподілу буде:

$$F(x)=\begin{cases}0, & x\le a\\ \dfrac{x-a}{b-a}, & a<x\le b\\ 1, & x>b\end{cases}. \qquad (59)$$

Графіки функцій $f(x)$, $F(x)$ мають вигляд (мал. 9):



Мал. 9

Числові характеристики випадкової величини, що має рівномірний розподіл на відрізку $[a,b]$ мають вигляд:

$$M(X)=\frac{a+b}{2};\;\; D(X)=\frac{(b-a)^2}{12};\;\; \sigma(X)=\frac{(b-a)\sqrt{3}}{6}. \qquad (60)$$

Дійсно. Щільність розподілу рівномірно розподіленої величини X задана рівністю (58), тому послідовно знаходимо $M(X)$, $D(X)$, $\sigma(X)$:

$$1)\; M(X)=\int_a^b x\frac{dx}{b-a}=\frac{1}{b-a}\frac{x^2}{2}\Big|_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{b+a}{2}.$$

Тобто математичне сподівання рівномірно розподіленої неперервної випадкової величини є середина заданого відрізка.

$$2)\; D(X)=\int_a^b x^2 f(x)dx-M^2(X)=\int_a^b\frac{x^2}{b-a}dx-\frac{(b+a)^2}{4}=\frac{1}{b-a}\frac{x^3}{3}\Big|_a^b-\frac{(b+a)^2}{4}=$$

$$=\frac{1}{b-a}\frac{b^3-a^3}{3}-\frac{b^2+2ab+a^2}{4}=\frac{a^2+ab+b^2}{3}-\frac{a^2+2ab+b^2}{4}=\frac{(b-a)^2}{12}$$

$$3)\,\sigma(X)=\sqrt{D(X)}=\sqrt{\frac{(b-a)^2}{12}}=\frac{b-a}{2\sqrt{3}}=\frac{(b-a)\sqrt{3}}{6},\quad (a<b).$$

Приклад 70. Неперервна випадкова величина X розподілена рівномірно на інтервалі $(0,1)$. Знайти ймовірність попадання X в інтервал $(\alpha,\beta)\subset(0,1)$.

Розв'язування. Використовуючи формулу (58), знайдемо щільність розподілу:

$$f(x)=\begin{cases}0, & x\le 0\\ 1, & 0<x\le 1\\ 0, & x>1\end{cases}.$$

Тоді,

$$P(\alpha<X<\beta)=\int_\alpha^\beta f(x)dx=\int_\alpha^\beta dx=x\Big|_\alpha^\beta=\beta-\alpha,$$

тобто шукана ймовірність дорівнює довжині відрізка $[\alpha,\beta]$.

Показниковий розподіл

Означення. ***Неперервна випадкова величина розподілена за показниковим законом з параметром*** $\lambda>0$***, якщо її щільність розподілу ймовірності має вигляд:***

$$f(x)=\begin{cases}0, & x<0\\ \lambda e^{-\lambda x}, & x\ge 0\end{cases} \qquad (61)$$

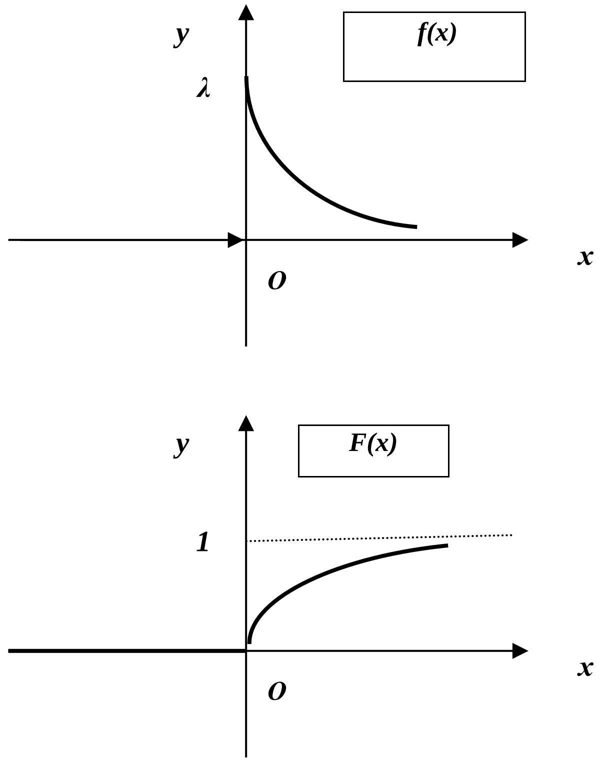
Відповідна функцію розподілу $F(X)$ буде:

$$F(x)=\int_{-\infty}^x f(t)dt=\int_0^x \lambda e^{-\lambda t}dt=\frac{\lambda e^{-\lambda t}}{-\lambda}\Big|_0^x=1-e^{-\lambda x}$$

Отже**,**

$$F(x)=\begin{cases}0 & при \quad x<0\\ 1-e^{-\lambda x} & при \quad x\ge 0\end{cases} \qquad (62)$$

Графіки функцій $f(x)$, $F(x)$ мають вигляд (мал. 10):



Мал. 10

Приклад 71. Нехай випадкова величина *X* розподілена за показниковим законом з параметром $\lambda=2$. Знайти ймовірність того, що випадкова величина *X* попаде в інтервал $(0.5, 2)$.

Розв'язування. Використовуючи властивість функції розподілу, знайдемо:

$$P(0.5<X<2)=F(2)-F(0.5)=e^{-1}-e^{-4}=0.3679-0.0733=0.2946 .$$

Числові характеристики випадкової величини *X*, що має показниковий розподіл з параметром λ, мають вигляд:

$$1)\, M(X)=\frac{1}{\lambda};\ 2) D(X)=\frac{1}{\lambda^2};\ 3)\sigma(X)=\frac{1}{\lambda}. \qquad (63)$$

Дійсно. Враховуючи вигляд функції щільності (61), знаходимо відповідно $M(X);D(X);\sigma(X)$:

$$1)\ M(X)=\int\limits_{-\infty}^{+\infty} xf(x)dx=\lambda\int\limits_{0}^{\infty} xe^{-\lambda x}dx=\lambda\lim_{b\to\infty}\left(-\frac{1}{\lambda}\right)\int\limits_{0}^{b} xd\left(e^{-\lambda x}\right)=$$

$$=-\lim_{b\to\infty}\left(xe^{-\lambda x}\Big|_0^b-\int\limits_0^b e^{-\lambda x}dx\right)=-\lim_{b\to\infty}\left(\frac{b}{e^{-\lambda b}}+\frac{1}{\lambda}e^{-\lambda x}\Big|_0^b\right)=$$

$$=-\lim_{b\to\infty}\left(\frac{b}{e^{-\lambda b}}-\frac{1}{\lambda}\right)=\frac{1}{\lambda}$$

;

$$2)\ D(X)=\int\limits_{-\infty}^{+\infty} x^2 f(x)dx-M^2(X)=-\int\limits_0^{\infty} x^2e^{-\lambda x}dx-\frac{1}{\lambda^2}=$$

$$=-\int\limits_0^{\infty} x^2 d\left(e^{-\lambda x}\right)-\frac{1}{\lambda^2}=-\lim_{b\to\infty}\left(x^2e^{-\lambda x}\Big|_0^b-\int\limits_0^b 2xe^{-\lambda x}dx\right)-\frac{1}{\lambda^2}=;$$

$$=-\lim_{b\to\infty}\left(\frac{b^2}{e^{-\lambda b}}-\frac{2}{\lambda}\cdot\lambda\int\limits_0^b xe^{-\lambda x}dx\right)-\frac{1}{\lambda^2}=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$$

$$3)\ \sigma(X)=\sqrt{D(X)}=\sqrt{\frac{1}{\lambda^2}}=\frac{1}{\lambda}.$$

Приклад 72. Знайти числові характеристики неперервної випадкової величини *X*, яка розподілена за показниковим законом:

$$f(x)=\begin{cases}0, & x<0\\ e^{-4x}, & x\geq 0\end{cases}.$$

Розв'язування. Випадкова величина *X* розподілена за показниковим законом з параметром $\lambda=4$. Згідно з рівностями (63), маємо:

$$M(X)=\sigma(X)=\frac{1}{\lambda}=\frac{1}{4}=0{,}25;\ D(X)=\frac{1}{\lambda^2}=\frac{1}{16}=0{,}0625.$$

Слід відмітити, що ймовірність попадання випадкової величини *X*, яка має показниковий розподіл в інтервал $(a,b)\subset(0,+\infty)$ буде рівна:

$$P(a\leq X<b)=\int\limits_a^b f(x)dx=\lambda\int\limits_a^b e^{-\lambda x}dx=e^{-\lambda a}-e^{-\lambda b}$$

Приклад 73. Відомо, що час безвідмовної роботи персонального комп`ютера розподіляється за показниковим законом $f(t)=0{,}01e^{-0{,}01t}$. Знайдемо ймовірність того, що персональний комп`ютер працюватиме не менше $t=200$ годин.

Розв'язування. Знаходимо:

$$P(T\geq 200)=1-P(T<200)=1-(1-e^{-0{,}01\cdot 200})\approx 0{,}1637.$$

Нормальний закон розподілу

Даний розподіл (***який називають законом Гаусса***) відіграє важливу роль в теорії ймовірностей та математичній статистиці і широко застосовується на практиці.

Означення. ***Випадкова величина X називається розподіленою нормально з параметрами*** a, σ***, якщо її щільність ймовірності має вигляд:***

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in (-\infty, +\infty), \qquad (64)$$

а функція розподілу її виражається невласним інтегралом:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(z-a)^2}{2\sigma^2}} dz \qquad (65)$$

Нормальний закон розподілу з параметрами a і σ позначають $N(a, \sigma)$.

З`ясуємо зміст параметрів a і σ нормального розподілу.

Покажемо, що для нормальної випадкової величини X математичне сподівання дорівнює параметру a, а середнє квадратичне відхилення дорівнює σ**.**

Дійсно, враховуючи щільність розподілу (64), знайдемо числові характеристики:

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

$$= \left| \begin{array}{l} заміна \ \ \dfrac{x-a}{\sigma} = z \Rightarrow x = \sigma z + a \Rightarrow dx = \sigma dz, \\ \qquad x \to -\infty \Rightarrow z \to -\infty, \ x \to +\infty \Rightarrow z \to +\infty \end{array} \right| =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + a) e^{-\frac{z^2}{2}} \sigma dz = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{z^2}{2}} dz + \frac{a\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz =$$

$$= \frac{a}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = a.$$

Тут використано рівності:

$$\int_{-\infty}^{\infty} z e^{-\frac{z^2}{2}} dz = 0 \ (\textit{для непарної функції}),$$

$$\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi} \ (\textit{інтеграл Пуассона});$$

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

$$=\left|\begin{array}{l}\dfrac{x-a}{\sigma}=z\Rightarrow x=\sigma z+a\Rightarrow dx=\sigma dz,\\ x\to-\infty\Rightarrow z\to-\infty,\ x\to+\infty\Rightarrow z\to+\infty\end{array}\right|=$$

$$=\frac{\sigma^2\cdot\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}z^2e^{-\frac{z^2}{2}}dz=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z\cdot ze^{-\frac{z^2}{2}}dz=-\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}zd\left(e^{-\frac{z^2}{2}}\right)=$$

$$=-\frac{2\sigma^2}{\sqrt{2\pi}}\int_{0}^{\infty}zd\left(e^{-\frac{z^2}{2}}\right)=-\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to\infty}\left(ze^{-\frac{z^2}{2}}\Big|_0^b-\int_0^b e^{-\frac{z^2}{2}}dz\right)=$$

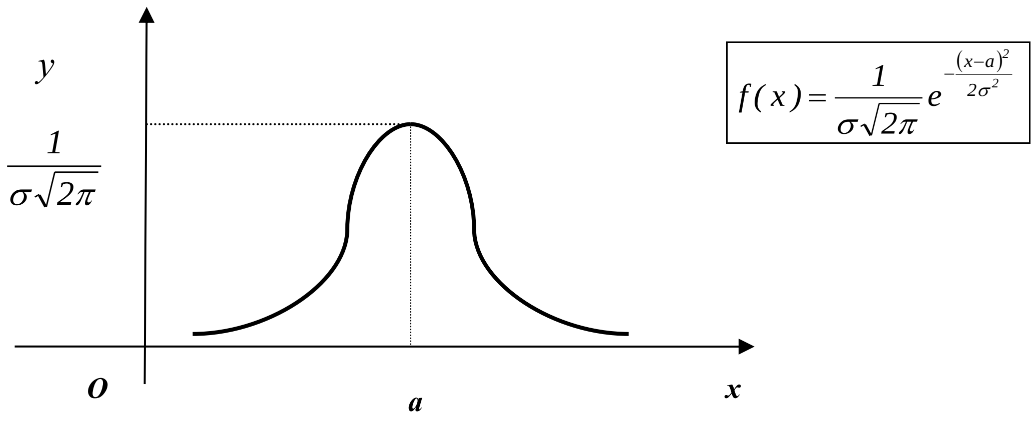
$$=-\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to\infty}\left(\underbrace{\frac{b}{e^{\frac{b^2}{2}}}}_{\to 0}-\int_0^b e^{-\frac{z^2}{2}}dz\right)=\frac{2\sigma^2}{\sqrt{2\pi}}\cdot\frac{\sqrt{2\pi}}{2}=\sigma^2.$$

Оскільки, $D(X)=\sigma^2$,

То

3) $\sigma(X)=\sqrt{D(X)}=\sqrt{\sigma^2}=\sigma$.

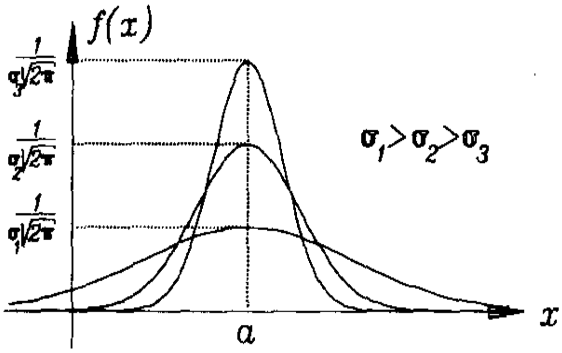
Графік щільності нормального розподілу представлено на мал. 11 ***(пропонуємо читачеві здійснити повне дослідження графіка щільності методами диференціального числення .***



Мал. 11

Зазначимо, що вісь *Ox* є горизонтальною асимптотою, а пряма $x = a$ - віссю симетрії графіка функції щільності. Досліджуючи $f'(x)$, знаходимо максимум функції $f(a)=\max\limits_{-\infty<x<\infty}f(x)=\frac{1}{\sigma\sqrt{2\pi}}$.

Крім того, зміна параметра *а* ***(математичного сподівання)*** не змінює форми кривої, а лише зсуває її вздовж вісі *Ox* вправо або вліво, в залежності від значення параметра *а* (див. мал. 12). Зміна ж параметра σ змінює форму кривої Гаусса. Якщо параметр σ ***(середньоквадратичне відхилення)*** збільшується, то $f_{\max} = f(a)$ зменшується, тобто крива стискується вздовж вісі Oy. Якщо ж σ спадає, то крива Гаусса розтягується у додатному напрямі вісі Oy. Зі збільшенням ***дисперсії*** σ^2 ймовірності значень, віддалених від центра розсіювання, збільшуються.



Мал. 12

Якщо випадкова величина X розподілена за нормальним законом з параметрами $a=0$ та $\sigma=1$, то такий розподіл називається ***нормованим нормальним або стандартним розподілом***. Функцією щільності у цьому випадку є ***функція Гаусса,*** графік якої симетричний відносно вісі Oy:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{66}$$

Інтеграл від функції $\varphi(x)$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz \tag{67}$$

називається ***функцією Лапласа***. Даний інтеграл не виражається в елементарних функціях і має широке практичне застосування.

Таблиці значень функцій $\varphi(x)$, $\Phi(x)$ наведені в додатках 1, 2.

Встановимо зв'язок між функцією розподілу нормальної випадкової величини $F(x)$ і функцією Лапласа $\Phi(x)$. Згідно з рівністю (65) маємо:

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

$$= \left| \begin{array}{ll} t = \dfrac{z-a}{\sigma}, & dz = \sigma dt \\ z \to -\infty \Rightarrow t \to -\infty, z = x \Rightarrow t = \dfrac{x-a}{\sigma} & \end{array} \right| =$$

$$=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\sigma dt=\frac{1}{\sqrt{2\pi}}\left[\int\limits_{-\infty}^{0} e^{-\frac{t^2}{2}}\sigma dt+\int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\sigma dt\right]=$$

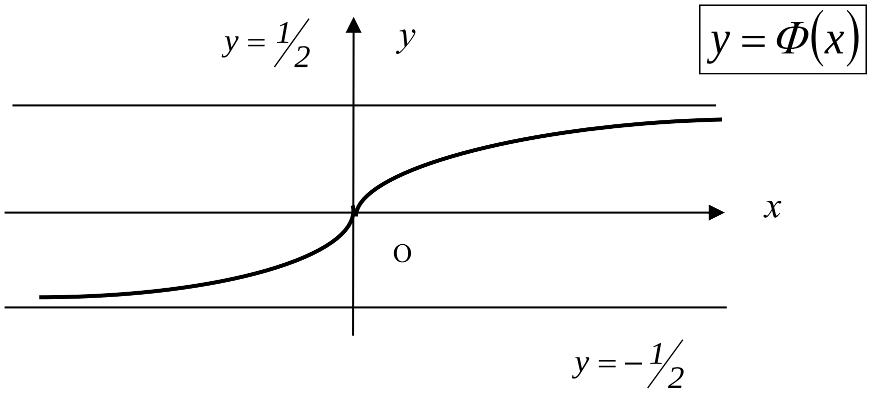
$$=\frac{1}{\sqrt{2\pi}}\left[\frac{\sqrt{2\pi}}{2}+\int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\sigma dt\right]=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\sigma dt$$

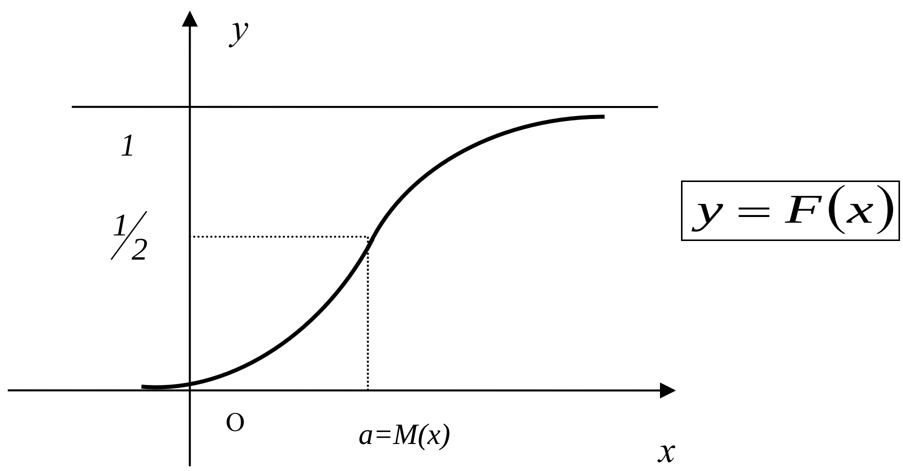
Отже,

$$F(x)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right), \tag{68}$$

де $\Phi\left(\frac{x-a}{\sigma}\right)=\Phi(z)$ - функція Лапласа.

Графіки функцій $\Phi(x)$, $F(x)$ зображено на мал. 13.





Мал. 13

($z=\frac{x-a}{\sigma}$, a - математичне сподівання випадкової величини X, σ - середнє квадратичне відхилення).

Означення. ***Число M_e називається медіаною розподілу, якщо виконується рівність:***

$$\int_{-\infty}^{M_e} f(x)dx = \int_{M_e}^{+\infty} f(x)dx = \frac{1}{2} \text{ або } P(X < M_e) = P(X > M_e) = \frac{1}{2}$$

Якщо випадкова величина X розподілена нормально, то її медіана M_e дорівнює a.

Ймовірність попадання нормально розподіленої випадкової величини X в інтервал (α, β)

Використовуючи формулу $P(\alpha < X < \beta) = F(\beta) - F(\alpha)$, а також рівність (68), одержимо ймовірність того, що нормально розподілена випадкова величина X належатиме інтервалу (α, β):

$$P(\alpha < X < \beta) = \frac{1}{2} + \Phi\left(\frac{\beta - a}{\sigma}\right) - \frac{1}{2} - \Phi\left(\frac{\alpha - a}{\sigma}\right) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \qquad (69)$$

Відмітимо, що ймовірність попадання нормально розподіленої випадкової величини X на проміжок довжини 2δ , симетричний відносно центру розсіювання $P(a - \delta < X < a + \delta)$ знаходиться за формулою:

$$P(a - \delta < X < a + \delta) = P\{|X - a| < \delta\} = 2\Phi\left(\frac{\delta}{\sigma}\right). \qquad (70)$$

Приклад 74. Випадкова величина X розподілена за нормальним законом. Математичне сподівання дорівнює 50, а середнє квадратичне відхилення – 10. Знайти ймовірність того, що випадкова величина X набуде значення з інтервалу *(10;50)*.

Розв'язування. За формулою (69) при $a = 50$, $\sigma = 10$ знайдемо шукану ймовірність, використовуючи непарність функції Лапласа $\Phi(x)$ та таблицю з додатку 2.

$$P(10 < X < 50) = \Phi\left(\frac{50 - 50}{10}\right) - \Phi\left(\frac{10 - 50}{10}\right) = \Phi(0) - \Phi(-4) \approx 0 + 0{,}4999 = 0{,}4999.$$

Приклад 75. Відбувається зважування деякої речовини без систематичних помилок. Випадкова помилка зважування розподілена за нормальним законом з середнім квадратичним відхиленням σ = 20 грам. Знайдемо ймовірність того, що зважування буде відбуватись з помилкою, яка не перевищує за абсолютною величиною 10 грам.

Розв'язування. Нехай випадкова величина X – помилка зважування. Оскільки систематичних помилок немає, то математичне сподівання $M(X)$ = 0. Застосуємо формулу (69). Покладемо в ній a = 0, σ = 20, δ = 10, знаходимо:

$$P\{|X| < 10\} = 2\Phi\left(\frac{10}{20}\right) = 2\Phi(0{,}5).$$

Згідно таблиці 2, знаходимо $\Phi(0{,}5) = 0{,}19146$. Отже, шукана ймовірність $P\{|X| < 10\} = 2 \cdot 0{,}19146 = 0{,}38292$.

Правило «трьох сигм»

Знайдемо ймовірність відхилення випадкової величини X від її математичного сподівання. В силу рівності (70), знаходимо при $\delta=\sigma$, $\delta=2\sigma$, $\delta=3\sigma$:

$$P\left(|X-a|<\sigma\right)=2\Phi\left(\frac{\sigma}{\sigma}\right)=2\Phi(1)\approx 2\cdot 0{,}3413=0{,}6826\ ,$$

$$P\left(|X-a|<2\sigma\right)=2\Phi\left(\frac{2\sigma}{\sigma}\right)=2\Phi(2)\approx 2\cdot 0{,}4772=0{,}9544\ ,$$

$$P\left(|X-a|<3\sigma\right)=2\Phi\left(\frac{3\sigma}{\sigma}\right)=2\Phi(3)\approx 2\cdot 0{,}49865=0{,}9971\approx 1\ . \qquad (71)$$

Отже, ***якщо випадкова величина X має нормальний розподіл, то ймовірність її відхилення від математичного сподівання за абсолютною величиною, яке менше трьох середньоквадратичних відхилень, близьке до одиниці.***

Іншими словами, ***подія*** $\left(|X-a|<3\sigma\right)$ ***є практично достовірною.*** В цьому і полягає «правило 3σ».

Зокрема, якщо неперервна випадкова величина розподілена за стандартним нормальним законом $N(0{,}1)$, то рівність (71) можна записати таким чином:

$$P\left(|X|<3\sigma\right)=2\Phi\left(\frac{3\sigma}{\sigma}\right)=2\Phi(3)\approx 1\ . \qquad (72)$$

Отримане правило означає, що результатами вимірювань з відхиленнями від математичного сподівання більше, ніж на три сигми, можна нехтувати. Такі міркування дозволяють суттєво підвищувати надійність даних, отриманих шляхом вимірювання. Важливо відмітити, що на практиці це правило використовують ще й таким чином: ***якщо закон розподілу випадкової величини X невідомий, але*** $|X-a|<3\sigma$***, то можна припускати, що випадкова величина X розподілена нормально.***

Завдання для самостійної роботи студентів

1. Випадкова величина X розподілена рівномірно в інтервалі *(1;10)*. Знайти: 1) щільність розподілу *X*, побудувати її графік; 2) функцію розподілу *X*, побудувати її графік; 3) числові характеристики *X*; 4) ймовірність попадання *X* в інтервал (-1; 4).

2. Потяг метро ходить з інтервалом 4 хвилини. Пасажир підходить до зупинки в деякий випадковий момент часу. Знайти ймовірність того, що цей пасажир чекатиме чергового потягу менше 2-х хвилин. Вважати, що час чекання є неперервна випадкова величина, розподілена рівномірно в інтервалі *(0;4)* (в інтервалі між двома сусідніми потягами).

3. Ціна шкали вимірювального приладу дорівнює 0,2 . Покази вимірювань округляють до найближчої цілої поділки. Знайти ймовірність того, що при

вимірюванні приладом буде допущено похибку: 1) що не перевищує *0,01*; 2) меншу ніж *0,04*; 3) більшу ніж *0,06*.

4. Тривалість знаходження клієнта банку в черзі має показниковий закон розподілу. Знайдіть імовірність того, що навмання обраний відвідувач буде знаходитися в черзі більше 3 хв., якщо середня тривалість знаходження в черзі 12хв.

5. Час ремонту приладу має показниковий закон розподілу з середнім значенням 5хв. Знайдіть імовірність того, що ремонт буде тривати від 5 до 10 хв.

6. Час обслуговування клієнта в банку має показниковий закон розподілу з середнім значенням 10 хв. Знайдіть імовірність того, що клієнта будуть обслуговувати час, більший за середній.

7. Час безвідмовної роботи банкомату є випадковою величиною з показниковим законом розподілу. Визначіть ймовірність безвідмовної роботи банкомату протягом місяця, якщо середній час безвідмовної роботи дорівнює шість місяців.

8. Тривалість телефонної розмови має показниковий закон розподілу. Середня тривалість розмови 3 хв. Визначіть імовірність того, що:

а) дві з трьох розмов будуть тривати більше середнього часу;

б) всі три розмови будуть тривати більше середнього часу.

9. Випадкова величина *X* розподілена за показниковим законом, заданим щільністю ймовірності :

$$f(x)=\begin{cases}0, & x<0\\ 2{,}4\cdot e^{-2{,}4x}, & x\geq 0\end{cases}.$$

Знайти: 1) функцію показникового розподілу; 2) числові характеристики *X*; 3) ймовірність того, що *X* набуде значення з інтервалу *(2;7)*.

10. Керуючий банком встановив, що число пропущених днів за хворобою за рік розподілено за нормальним законом із математичним сподіванням 67 днів і середнім квадратичним відхиленням 10 днів. Яка ймовірність того, що число пропущених днів за хворобою в цьому році:

а) від 30 до 50;

б) менше 30;

в) більше 60.

11. Вага упаковок розфасованого цукру розподілена за нормальним законом із математичним сподіванням 1 кг і середнім квадратичним відхиленням 0,03 кг. Яка ймовірність, що навмання взята упаковка важить:

а) від 1,1 до 1,2 кг;

б) від 0,9 до 1,15 кг;

в) більше 1,1 кг;

г) менше 1,1 кг.

12. Фірма реалізує товар через мережу магазинів. Ціна товару *X* має нормальний закон розподілу з математичним сподіванням (проектною ціною) 50 грн. Імовірність того, що ціна товару, що реалізується, не менше 32 грн і не більше 68 грн дорівнює 0,96. Знайдіть імовірність того, що ціна товару в навмання обраному магазині: а) менше 45 грн; б) більше 45 грн.

13. Випадкова величина має нормальний закон розподілу з середнім квадратичним відхиленням 12. Знайдіть імовірність того, що модуль відхилення випадкової величини від її математичного сподівання буде більшим за 25.

14. Випадкова величина має нормальний закон розподілу з середнім квадратичним відхиленням 5. Знайдіть ймовірність того, що модуль відхилення випадкової величини від її математичного сподівання буде меншим 4.

15. Прибуток відділення банку має нормальний закон розподілу з математичним сподіванням 120000 грн. в місяць і середнім квадратичним відхиленням 10000 грн. Визначити, в яких межах знаходиться прибуток банку з надійністю 93%.

16. Розглядається можливість придбання акцій двох видів. Їх прибутковості розподіляються за нормальним законом з однаковим математичним сподіванням 15% і середнім квадратичним відхиленням 2,5% та 3% відповідно. Визначіть інтервал прибутковості акції кожного виду з ймовірністю 0,95.

17. Випадкова величина *X* розподілена за нормальним законом з параметрами $a = 3$, $\sigma = 1.2$. 1) Знайти функцію щільності розподілу випадкової величини *X*; 2) визначити ймовірність того, що значення випадкової величини *X* будуть знаходитися в інтервалі *(2;5)*.

18. Випадкова величина *X* розподілена за нормальним законом з математичним сподіванням $a = 15$ і середнім квадратичним відхиленням $\sigma = 5$. Визначити інтервал, в який потрапить з ймовірністю 0,9967 в результаті випробувань випадкова величина X .

19. Сума вкладу у відділенні банку є випадковою величиною, розподіленою за нормальним законом з математичним сподіванням 5000 грн. і середнім квадратичним відхиленням 400 грн. Яку максимальну суму вкладу можна гарантувати з імовірністю 0,95?

Глава 2. Граничні теореми теорії ймовірностей

Практика застосування теорії ймовірностей показала, що при спостереженні великої кількості випадкових величин їх середній результат перестає бути випадковим і може бути передбачений з великим ступенем визначеності. Це дозволило сформулювати ряд теорем, у кожній з яких при тих, або інших умовах встановлюється факт наближення середніх характеристик при великому числі випробувань до деяких визначених сталих. Дані результати широко застосовують у техниці, економіці, фінансах, банківській справі, соціології та ін. Вони є теоретичною основою математичної статистики, економетрики, фінансової та актуарної математики. Граничні теореми встановлюють закони розподілу сум великої кількості незалежних випадкових величин. При їх доведенні, а також при розв'язуванні багатьох практичних задач, використовується нерівність Чебишева.

§1. Нерівність Чебишева

Лема ***(Чебишева).*** Для випадкової величини X, що має скінчені математичне сподівання $M(X)$ та дисперсію $D(X)$, при довільному $\varepsilon > 0$ виконується нерівність:

$$P\left(\left|X - M(X)\right| \geq \varepsilon\right) \leq \frac{D(X)}{\varepsilon^2} . \qquad (73)$$

Доведення. Доведемо цю нерівність для дискретної випадкової величини.

Розглянемо випадкову величину X з законом розподілу:

X	x_1	x_2	$\cdots$	x_n
p	p_1	p_2	$\cdots$	p_n

Події $\left|X - M(X)\right| < \varepsilon$ і $\left|X - M(X)\right| \geq \varepsilon$ є протилежними, а отже сума їх ймовірностей рівна одиниці:

$$P\left(\left|X - M(X)\right| < \varepsilon\right) + P\left(\left|X - M(X)\right| \geq \varepsilon\right) = 1 .$$

Враховуючи означення дисперсії $D(X)$, знаходимо:

$$D(X) = M\left(X - M(X)\right)^2 = \sum_{k=1}^{n} p_k \left(x_k - M(X)\right)^2 =$$

$$= \sum_{k:\left|x_k - M(X)\right| < \varepsilon} p_k \left(x_k - M(X)\right)^2 + \sum_{k:\left|x_k - M(X)\right| \geq \varepsilon} p_k \left(x_k - M(X)\right)^2 \geq$$

$$\geq \sum_{k:\left|x_k - M(X)\right| \geq \varepsilon} p_k \left(x_k - M(X)\right)^2 \geq \sum_{k:\left|x_k - M(X)\right| \geq \varepsilon} p_k \varepsilon^2 =$$

$$= \varepsilon^2 \sum_{k:\left|x_k - M(X)\right| \geq \varepsilon} p_k = \varepsilon^2 P\left(\left|X - M(X)\right| \geq \varepsilon\right).$$

Звідси маємо нерівність (73).

Зауважимо, що нерівність Чебишева часто застосовується і в іншій формі:

$$P\left(\left|X - M(X)\right| < \varepsilon\right) \geq 1 - \frac{D(X)}{\varepsilon^2}\quad. \qquad (74)$$

Відмітимо також, що дане твердження справедливе як для дискретних, так і для неперервних випадкових величин *X*. Нерівність Чебишева дозволяє оцінювати ймовірності відхилень значень випадкової величини від своїх математичних сподівань. Нерівність Чебишева в формі (73) встановлює верхню межу ймовірності події, а у формі (74) – нижню межу.

За допомогою нерівності (73) оцінимо похибку наближення вимірювань.

Приклад 76. Припустимо, що проводиться n незалежних вимірів деякої величини a. Похибки вимірювань $\delta_1, \delta_2, \ldots, \delta_n$ є випадковими величинами. Вважаємо також, що $M(\delta_k) = 0, k = \overline{1, n}$, тобто, що відсутня систематична похибка приладів вимірювання. Нехай $D(\delta_k) = b^2, k = \overline{1, n},$ тобто прилади забезпечують певну точність вимірів.

Розв'язування. За значення невідомої величини a приймають середнє арифметичне результатів вимірів. Тоді похибка при визначенні числа a буде дорівнювати

$$\Delta_n = \frac{1}{n}\sum_{k=1}^{n}\delta_k,$$

Звідси дисперсія $D(\Delta_n) = \frac{1}{n^2}\sum_{k=1}^{n} D(\delta_k) = \frac{nb^2}{n^2} = \frac{b^2}{n},$ а математичне сподівання -

$$M(\Delta_n) = \frac{1}{n}\sum_{k=1}^{n} M(\delta_k) = 0.$$

Нехай похибка Δ_n не перевищує деякого $\varepsilon > 0$ з ймовірністю ω, наприклад при ω=0,99:

$$P\left(\left|\Delta_n\right| < \varepsilon\right) \geq 0{,}99. \qquad (75)$$

Тоді згідно з нерівністю Чебишева (74) маємо:

$$P\left(\left|\Delta_n\right| < \varepsilon\right) \geq 1 - \frac{D(\Delta_n)}{\varepsilon^2} = 1 - \frac{b^2}{n\varepsilon^2}.$$

Таким чином, нерівність (75) буде виконуватися за умови:

$$1 - \frac{b^2}{n\varepsilon^2} \geq 0{,}99 \Leftrightarrow \frac{b^2}{n\varepsilon^2} \leq 0{,}01, \text{ або } n \geq \frac{100b^2}{\varepsilon^2}.$$

Отже, отримана оцінка числа вимірювань, необхідних для досягнення заданої точності.

§2. *Закон великих чисел*

Теорема (*закон великих чисел у формі Чебишева*). *Якщо випадкові величини $X_1, X_2, \dots, X_n, \dots$ попарно незалежні, мають скінченні математичні сподівання і дисперсії, і виконується співвідношення:*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} D(X_k) = 0, \tag{76}$$

то для довільного $\varepsilon > 0$ маємо рівність:

$$\lim_{n\to\infty} P\left(\left| \frac{1}{n} \sum_{k=1}^{n} X_k - \frac{1}{n} \sum_{k=1}^{n} M(X_k) \right| < \varepsilon \right) = 1. \tag{77}$$

Доведення. Позначимо $\overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$. Тоді (77) еквівалентне тому, що для будь-якого $\varepsilon > 0$ справедлива рівність:

$$\lim_{n\to\infty} P\left(\left| \overline{X}_n - M\overline{X}_n \right| \geq \varepsilon \right) = 0. \tag{78}$$

Випадкові величини $X_1, X_2, \dots, X_n$ попарно незалежні, тому

$$D(\overline{X}_n) = \frac{1}{n^2} \sum_{k=1}^{n} D(X_k). \tag{79}$$

З застосуванням нерівності Чебишева і рівності (79), знаходимо:

$$P\left(\left| \overline{X}_n - M\overline{X}_n \right| \geq \varepsilon \right) \leq \frac{D(\overline{X}_n)}{\varepsilon^2} = \frac{\frac{1}{n^2} \sum_{k=1}^{n} D(X_k)}{\varepsilon^2}. \tag{80}$$

Далі, згідно з різностями (79) і (76), дістанемо (78).

Дійсно,

$$\lim_{n\to\infty} P\left(\left| \overline{X}_n - M\overline{X}_n \right| \geq \varepsilon \right) \leq \lim_{n\to\infty} \frac{\frac{1}{n^2} \sum_{k=1}^{n} D(X_k)}{\varepsilon^2} = 0.$$

Твердження доведено.

Зауваження 1. Теорема Чебишева також справедлива, якщо випадкові величини $X_1, X_2, \dots, X_n, \dots$ попарно незалежні, мають скінченні математичні сподівання і дисперсії, обмежені в сукупності, тобто:

$$D(X_k) \leq C, k = \overline{1, \infty},$$

де C - стала величина.

Рівність (77) означає, що середнє арифметичне значень випадкових величин, коли кількість доданків нескінченно зростає збігається до середнього арифметичного їх математичних сподівань. Збіжність, яка зазначена у співвідношенні (77) називається ***збіжністю за ймовірністю.***

Зауваження 2. Для застосування теореми Чебишева на практиці, її формулюють наступним чином: ***якщо дисперсії незалежних випадкових величин обмежені в сукупності, то для великих n має місце наближена рівність:***

$$\frac{1}{n}\sum_{k=1}^{n} X_k \approx \frac{1}{n}\sum_{k=1}^{n} M(X_k). \tag{81}$$

Приклад 77. Дисперсія кожної з 5000 незалежних випадкових величин не перевищує 3. Оцінити ймовірність того, що відхилення середнього арифметичного цих випадкових величин від середнього арифметичного їх математичних сподівань не перевищить за абсолютною величиною числа $\varepsilon = 0{,}2$

Розв'язування. За умовою $n = 5000$, $\varepsilon = 0{,}2$, $C = 3$. Після підстановки цих значень у вираз (80) одержимо шукану оцінку

$$P\left(\left|\frac{X_1 + X_2 + \ldots + X_{5000}}{5000} - \frac{M(X_1) + M(X_2) + \ldots + M(X_{5000})}{5000}\right| < 0{,}2\right) \geq$$

$$\geq 1 - \frac{3}{5000 \cdot 0{,}04} = 0{,}985$$

Приклад 78. Оцінимо кількість доданків, яку слід взяти, щоб з ймовірністю (надійністю) 95% і точністю 0,01 виконувалася рівність (81).

Розв'язування. За умовою $\varepsilon = 0{,}01$. Для гарантування надійності 95%, оцінимо ймовірність події, протилежної до тієї, що розглядалася у (80):

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} M(X_k)\right| < \varepsilon\right) \geq 1 - \frac{D(X_k)}{n^2\varepsilon^2} = 1 - \frac{nC}{n^2\varepsilon^2} = 1 - \frac{C}{n\varepsilon^2}, \tag{82}$$

де $D(X_k) \leq C$.

Звідки знаходимо:

$$\frac{C}{\varepsilon^2 n} \leq 0{,}05 \Rightarrow n \geq \frac{C}{0{,}0001 \cdot 0{,}05} = 200000C.$$

Даний приклад показує, що для досягнення потрібної точності слід брати велику кількість доданків.

Наслідок (*теорема Чебишева про середнє арифметичне*). ***Для однаково розподілених, попарно незалежних випадкових величин*** $X_1, X_2, \ldots, X_n, \ldots$ ***з обмеженими в сукупності дисперсіями при довільному*** $\varepsilon > 0$ ***виконується рівність:***

$$\lim_{n \to \infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a\right| < \varepsilon\right) = 1, \text{ де } a = M(X_k), k = \overline{1, \infty}. \tag{83}$$

Цей наслідок з теореми Чебишева є обгрунтуванням правила «середнього арифметичного» в теорії вимірів.

Також класичним застосуванням закону великих чисел є теорема Бернуллі.

Теорема ***(теорема Бернуллі). Якщо*** k ***- кількість успіхів у n випробуваннях Бернуллі, а*** p $(0 < p < 1)$ ***- ймовірність успіху у кожному випробуванні, то для довільного*** $\varepsilon > 0$ ***виконується рівність:***

$$\lim_{n\to\infty} P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = 1. \qquad (84)$$

Схема Бернуллі є математичною моделлю серій випробувань, що повторюються за однакових умов. В кожному випробуванні з ймовірністю p настає подія *A*, яку називають «успіхом». Згідно теореми Бернуллі, частота k/n настання події *A* наближається до її ймовірності p . Цей факт підтверджується і на практиці.

Приклад 79. ***(Застосування нерівності Чебишева).*** Підприємець бере в банку кредит під 10% річних для впровадження нових технологій. При цьому експерти банку оцінюють, що ризик, пов'язаний з коливанням сподіваних прибутків підприємця становить 5%. Необхідно з ймовірністю $8/9 \approx 0{,}89$ оцінити (у %) рівень сподіваних прибутків для уникнення банкрутства.

Розв'язування. Нехай R – прибутковість діяльності підприємця є випадковою величиною. Тоді ймовірність банкрутства підприємця, тобто подія *(R<r),* згідно нерівності Чебишева (73) оцінюється так:

$$P(R < r) = P(-(R-m) > m-r) \le P(|R-m| > m-r) \le \frac{D(R-m)}{(m-r)^2},$$

де m – математичне сподівання випадкової величини R, r – відсоток по кредиту. При цьому враховується, що середня прибутковість діяльності підприємця m повинна бути вищою за відсоток по кредиту. Тобто: $m - r > 0$.

Для того, щоб ймовірність збанкрутувати була не більше ніж *1/9,* достатньо виконання умови, що випливає з правила 3-х сігм:

$$m - r > 3\sigma,$$

або

$$m > r + 3\sigma.$$

Дійсно, ймовірність банкрутства $P(R < r)$ буде менше *1/9* , якщо $\frac{D(R-m)}{(m-r)^2} < \frac{1}{9}$. Тобто, повинно виконуватись співвідношення:

$$(m-r)^2 > 9D(R-m) = 9\sigma^2.$$

Далі, згідно умови задачі, можемо записати:

$$m > 10\% + 3 \cdot 5\% = 25\%.$$

Тобто рівень (норма) сподіваних прибутків має бути не меншим ніж 25%.

Приклад 80. Оцінити ймовірність того, що випадкова величина X відхилиться від свого математичного сподівання 1) не менше ніж на два середніх квадратичних відхилення; 2) менше ніж на чотири середніх квадратичних відхилення.

Розв'язування. 1) Підставивши в нерівність Чебишева (73) значення $\varepsilon = 2\sigma$, одержимо:

$$P\left(\left|X-M(X)\right|\geq 2\sigma\right)\leq \frac{\sigma^2}{4\sigma^2}=\frac{1}{4}.$$

2) Підставивши в нерівність (74) значення $\varepsilon=4\sigma$, маємо:

$$P\left(\left|X-M(X)\right|<4\sigma\right)\geq 1-\frac{\sigma^2}{16\sigma^2}=1-\frac{1}{16}=\frac{15}{16}.$$

Приклад 81. Оцінити ймовірність того, що випадкова величина X відхилиться від свого математичного сподівання 1) менше ніж на три середніх квадратичних відхилення; 2) не менше ніж на чотири середніх квадратичних відхилення**.**

Розв'язування. Аналогічно розв‘язуванню прикладу 80, маємо:

$$1)\ P\left(\left|X-M(X)\right|<3\sigma\right)\geq 1-\frac{\sigma^2}{9\sigma^2}=\frac{8}{9};$$

$$2)\ P\left(\left|X-M(X)\right|\geq 4\sigma\right)\leq \frac{\sigma^2}{16\sigma^2}=\frac{1}{4}.$$

Приклад 82. Ймовірність несплати іпотечної позики, виданої банком становить 0,3. Кредитний інспектор перевіряє 600 позик. Оцінити ймовірність того, що число позик з порушенням термінів сплати буде: 1) меншим 40; 2) не меншим 40.

Розв'язування. 1) Нехай дискретна випадкова величина X – число позик з порушенням термінів сплати. Припускаємо, що позики незалежні. Отже, X розподілена за біномним законом. Тому

$$M(X)=np=600\cdot 0{,}3=180;\ D(X)=npq=600\cdot 0{,}3\cdot 0{,}7=126.$$

Скористаємось нерівністю Чебишева: $P\left(\left|X-M(X)\right|<\varepsilon\right)\geq 1-\dfrac{D(X)}{\varepsilon^2}$.

Оцінимо шукану ймовірність при $M(X)=180, D(X)=126, \varepsilon=40$:

$$P\left(\left|X-180\right|<40\right)\geq 1-\frac{126}{1600}\approx 1-0{,}921=0{,}921.$$

2) Події $|X-180|<40$ та $|X-180|\geq 40$ є протилежними, тому сума їх ймовірностей дорівнює одиниці.

Отже маємо:

$$P\left(\left|X-180\right|\geq 40\right)\leq 1-0{,}921=0{,}079.$$

Зауважимо, що в умовах кризи неповернення іпотечних позик стає залежними подіями і застосування схеми Бернуллі є неможливим.

§3. Центральна гранична теорема

Центральна гранична теорема встановлює закони розподілу сум великої кількості незалежних випадкових величин. Цей результат вперше був розглянутий в роботах Ляпунова А.М. і Маркова А.А.

Теорема ***(Ляпунова). Якщо для послідовності взаємно незалежних випадкових величин*** $\eta_1, \eta_2, \ldots \eta_n, \ldots$ ***існує таке*** $\delta > 0$***, що***

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} M|\eta_i - M\eta_i|^{2+\delta}}{\left(\sqrt{\sum_{i=1}^{n} D\eta_i}\right)^{2+\delta}} = 0, \tag{85}$$

то рівномірно по x ***виконується співвідношення:***

$$\lim_{n\to\infty} P\left\{ \frac{\sum_{i=1}^{n} \eta_i - \sum_{i=1}^{n} M\eta_i}{\sqrt{\sum_{i=1}^{n} D\eta_i}} < x \right\} = \Phi(x).$$

Зауважимо, що умова теореми Ляпунова (85) виконується, зокрема, якщо $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ незалежні та однаково розподілені, тобто їх функції розподілу однакові. Це означає, що їх математичні сподівання і дисперсії рівні, тобто має місце твердження.

Теорема. ***Для випадкових величин*** $\eta_1, \eta_2, \ldots, \eta_n, \ldots$***, які однаково розподілені, тобто:*** $M\eta_i = a, D\eta_i = \sigma^2, i = 1,2\ldots$ ***, виконується гранична рівність:***

$$\lim_{n\to\infty} P\left\{ \frac{\sum_{i=1}^{n} \eta_i - na}{\sigma\sqrt{n}} < x \right\} = \Phi(x).$$

Таким чином, зміст центральної граничної теореми полягає у тому, що сума великої кількості незалежних випадкових чинників при певному їх нормуванні має граничний стандартний нормальний розподіл, тобто зникає залежність від розподілів її окремих доданків.

Приклад 83. У відділення банку щодня заходить N клієнтів, кожний із яких робить вклад S з ймовірністю 0.5 і знімає суму S_1 з ймовірністю 0.5. Яку суму треба мати у відділенні банку на початок робочого дня, щоб банк міг здійснювати свої зобов'язання перед клієнтами з надійністю $g=96\%$ і $g=100\%$? Розв`язати задачу при $N=100$, $S=1000$ *грн.*, $S_1= 1000$ *грн.*

Розв'язування. 1) при $g=100\%$ потрібно мати у відділенні на початок дня суму $NS_1 =100000$ *грн.*, оскільки може статися, що усі клієнти знімуть з своїх рахунків суму S_1.

2) при надійності $g=96\%$ використаємо схему Бернуллі, де N кількість незалежних випробувань, $p=0.5$ – ймовірність успіху (зняття суми з рахунку), а $q= 0.5$ – ймовірність неуспіху (вклад на рахунок). Нехай m – кількість успіхів в серії довжини N . Тоді, як відомо: $Mm=Np$, $Dm=Npq$, тобто у нашому випадку

$Mm=N/2$, $Dm=N/4$. Знайдемо з ймовірністю g довжину інтервалу:

$$P\{0 \le m \le b\} = g \quad .$$

Згідно інтегральної теореми Муавра-Лапласа запишемо:

$$g = P\{0 \le m \le b\} \approx \Phi(\frac{b-\frac{N}{2}}{\sqrt{\frac{N}{4}}}) - \Phi(\frac{0-\frac{N}{2}}{\sqrt{\frac{N}{4}}}) = \Phi(\frac{2b-N}{\sqrt{N}}) + \Phi(\sqrt{N}) \quad .$$

При $N=100$ одержимо, $0.96 = \Phi(\frac{2b-100}{\sqrt{100}}) + \Phi(\sqrt{100})$, звідки

$$0.46 = \Phi(\frac{2b-100}{10}) \quad \text{і} \quad \frac{2b-100}{10} = 1.75,\ b=59.$$

Тобто, не більше 59 клієнтів банку (з ймовірністю 96%) знімуть суму S_1 , а отже вкладуть кошти – 41 клієнт. В цьому випадку баланс банку за добу складе :

$$B = -59S_1 + 41S = -18S,$$ (за умовою задачі $S = S_1$).

Отже, суму 18000 грн. і потрібно мати банку на початок робочого дня.

Відмітимо, що при високій надійності роботи (*96%)* вдалося зменшити суму грошового забезпечення відділення банку більше ніж у п’ять разів.

Приклад 84. При порівнянні кредитних відсоткових ставок виявлено, що по 1000 кредитах середня ставка на 2% більше, ніж у минулому році ($\sigma = 8\%$). Чи є це відхилення випадковим?

Розв'язування. Нехай випадкова величина – ставка по k –му кредиту має математичне сподівання та дисперсію $M\eta_k$, $D\eta_k \le 64$. (Вважаємо, що кредитні відсоткові ставки змінюються незалежно).

Використовуючи закон великих чисел Чебишева (77), оцінимо ймовірність середнього відхилення ставок більше ніж на 2%:

$$P\left\{\frac{1}{1000}\sum_{k=1}^{1000}(\eta_k - M\eta_k) \ge 2\right\} \le \frac{8^2}{1000 \cdot 2^2} \approx 0.016,$$

що складає 1,6%. Тобто ця подія малоймовірна. Іншими словами збільшення кредитних ставок на 2% у цьому році є невипадковим.

Завдання для самостійної роботи студентів

1. Дисперсія кожної з 2000 незалежних випадкових величин не перевищує 2. З якою ймовірністю можна стверджувати, що відхилення середнього арифметичного цих величин від середнього арифметичного їх математичних сподівань за абсолютною величиною менше ніж 0,07?

2. Імовірність появи події A в кожному з 180 випробувань дорівнює 0,25. Використовуючи нерівність Чебишева, оцінити імовірність того, що відхилення числа появи події A від математичного сподівання буде більшим за 50.

3. Дисперсія кожної з незалежних випадкових величин не перевищує 6. Визначити з ймовірністю, не меншою ніж 0,95, число таких величин, при якому відхилення їх середнього арифметичного від середнього арифметичного їх математичних сподівань буде меншим числа 0,3.

4. У результаті моніторингу за споживанням електроенергії підприємства встановлено, що середнє значення добової її потреби дорівнює 250 квт. Оцінити ймовірність того, що добова потреба підприємства в електроенергії буде не меншою 300 квт.

5. Тривалість депозиту вкладників банку є випадковою величиною, середнє значення якої дорівнює 180 днів. Дисперсія цієї величини дорівнює 100. Оцінити ймовірність того, що: 1) відхилення тривалості депозиту від його середнього значення за абсолютною величиною буде меншим ніж 5 днів; 2) тривалість депозиту знаходиться в межах від 150 до 190 днів.

6. Відомо, що 70% усієї продукції підприємства виробляється першим сортом. Оцінити ймовірність того, що число виробів першого сорту серед 40000 виготовлених буде відрізнятися за абсолютною величиною від математичного сподівання цього числа менше ніж на 100 штук.

7. Вибірковим способом необхідно визначити середню вагу кавових зерен. Скільки потрібно обстежити таких зерен, щоб з ймовірністю, не меншою 0,95, можна було б стверджувати, що середня вага відібраних кавових зерен буде відрізнятися від математичного сподівання менше ніж на 0,01г? Відомо, що середнє квадратичне відхилення ваги зерен не перевищує 0,03г.

8. Прилад складається з 20 незалежно працюючих елементів. Імовірність відмовлення кожного елемента за добу дорівнює $0{,}05$. За допомогою нерівності Чебишева оцінити ймовірність того, що абсолютна величина різниці між числом приладів, що відмовили, і середнім числом відмовлень за добу виявиться: 1) більше 3, 2) менше 5.

9. Для визначення середньої заробітної плати працівників малих підприємств в податковій інспекції вибірковим способом обстежили 700 підприємств. За отриманими даними обчислили середню заробітну плату (вибіркову середню зарплату). Відомо, що середні квадратичні відхилення заробітної плати працівників малих підприємств не перевищують 250 грн. Яке граничне відхилення вибіркової середньої від середнього значення, заробітної плати можна гарантувати з ймовірністю, не меншою: 1) 0,87;
2) 0, 95?

Глава 3. Системи випадкових величин

§1. Закони розподілу двовимірних випадкових величин

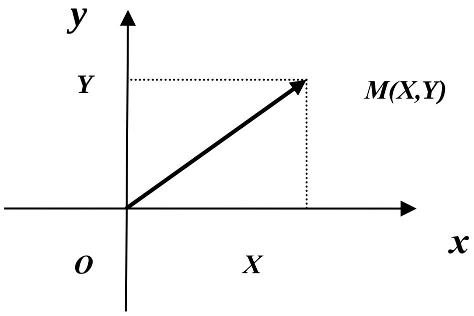
В попередніх параграфах розглядалися одновимірні випадкові величини. В застосуваннях теорії ймовірностей також вивчають випадкові величини, можливі значення яких визначені двома, трьома,.., п числами. Такі випадкові величини називаються відповідно двовимірними, трьохвимірними,.., n-вимірними випадковими величинами.

Приклад 85. Річні показники діяльності комерційного банку є випадковими. Так активи, пасиви, прибуток, резерви та інші показники належать до випадкових величин, оскільки їх значення не можна передбачити наперед.

Якщо контролюються лише показники активів X та пасивів Y, то маємо двовимірну випадкову величину (X, Y).

Означення. ***Сукупність випадкових величин (X,Y), які розглядаються одночасно, називається системою двох випадкових величин.***

Геометрично двовимірну випадкову величину можна представити, як випадкову точку $M(X,Y)$ на площині Oxy, або як випадковий вектор $\overline{OM}(X;Y)$ (мал. 14).



Мал. 14

Двовимірні величини, як і одновимірні, бувають дискретними і неперервними. Аналогічно визначається система п *випадкових величин, як випадковий* п-*мірний вектор* $(X_1, X_2,\dots,X_n)$.

Означення. ***Двовимірна випадкова величина (X,Y) називається дискретною, якщо її компоненти X та Y є дискретними випадковими величинами.***

Означення. ***Двовимірна випадкова величина (X,Y) називається неперервною, якщо її компоненти X та Y є неперервними випадковими величинами.***

Розподіл дискретної двовимірної випадкової величини

Закон розподілу дискретної двовимірної випадкової величини задають таблицею:

X \ Y	x_1	x_2	$\dots$	x_n
y_1	p_{11}	p_{11}	$\dots$	p_{11}
y_2	p_{21}	p_{21}	$\dots$	p_{21}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_m	p_{m1}	p_{m1}	$\dots$	p_{m1}

де (x_i, y_j) $(i = 1,..., n,\ \ j = 1,..., m)$ – значення двовимірної випадкової величини, $p_{ij} = P(x_i, y_j)$ – відповідні їм ймовірності. У першому рядку таблиці записані усі можливі значення компоненти *X*, а в першому стовпчику – *Y*. Оскільки, події в таблиці несумісні і утворюють повну групу, то справедлива рівність:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} = 1. \tag{86}$$

Закон розподілу двовимірної випадкової величини дозволяє знайти закони розподілу кожної з компонент. Дійсно, події $(x_i, y_1), (x_i, y_2),..., (x_i, y_m)$ несумісні, тому ймовірність $P(x_i)$ того, що випадкова величина *X* набуде значення x_i, за теоремою додавання ймовірностей буде дорівнювати

$$P(x_i) = P(X = x_i) = P(x_i, y_1) + P(x_i, y_2) + ... + P(x_i, y_m) = \sum_{j=1}^{n} p_{ij}\ ,\ i = \overline{1, n}\,,$$

тобто сумі ймовірностей, розміщених в i -му стовпчику.

Аналогічно можна записати закон розподілу для компоненти *Y*: $P(y_i) = P(Y = y_i) = \sum_{i=1}^{n} p_{ij}\,,\ j = \overline{1, m},$ тобто рівна сумі ймовірностей i -го рядка даної таблиці.

Обернене твердження в загальному випадку не має місця, оскільки знаючи розподіл кожної координати двовимірного вектора не можна в загальному випадку поновити їх сукупний розподіл.

Приклад 86. Знайти закони розподілу компонент двовимірної випадкової величини *(X, Y)*, закон розподілу якої задано таблицею:

X \ Y	-2	3	7
4	0,1	0,2	0,3
9	0,2	0,05	0,15

Розв'язування. Закони розподілу компонент X і Y:

X	-2	3	7
p	0,3	0,25	0,45

Y	4	9
p	0,6	0,4

Оскільки: $P(-2) = 0{,}1 + 0{,}2 = 0{,}3$;

$P(3) = 0{,}2 + 0{,}05 = 0{,}25$; $P(7) = 0{,}3 + 0{,}15 = 0{,}45$;

$P(4) = 0{,}1 + 0{,}2 + 0{,}3 = 0{,}6$; $P(9) = 0{,}2 + 0{,}05 + 0{,}15 = 0{,}4$.

Тобто, знаючи сукупний розподіл двовимірної випадкової величини , можна визначити закони розподілів її складових.

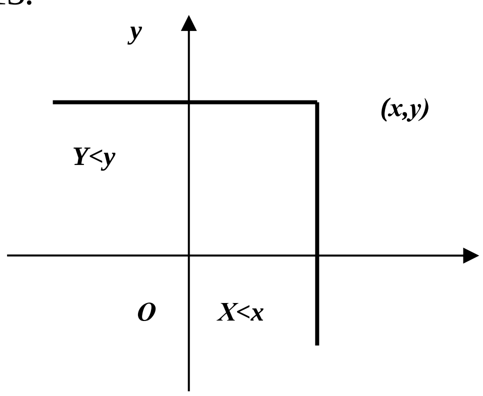
Багатовимірні випадкові величини, як і одновимірні, зручно характеризувати функціями розподілу.

Функція розподілу двовимірної випадкової величини та її властивості

Означення. ***Функцією розподілу двовимірної випадкової величини (X,Y) називається функція двох змінних*** $F(x,y)$***, яка визначає для кожної пари чисел (x,y) ймовірність виконання нерівностей X<x, Y<y , тобто***

$$F(x,y) = P(X < x, Y < y). \tag{87}$$

Геометрично рівність (87) означає: функція $F(x,y)$ є ймовірністю того, що випадкова точка $M(X,\ Y)$ попадає у нескінченний південно-західний сектор з вершиною (*x*,*y*), див. мал. 15.



Мал. 15

Властивості функції розподілу $F(x,y)$

Властивості функції розподілу двовимірної випадкової величини подібні властивостям функції розподілу одновимірної випадкової величини:

1. $0 \le F(x,y) \le 1$ (за означенням).
2. $F(x,y)$ - неспадна функція за кожним з аргументів, тобто:

$F(x_2, y) \ge F(x_1, y) \quad \forall x_2 > x_1$; $F(x, y_2) \ge F(x, y_1) \ \forall\ y_2 > y_1$

3. Для функції $F(x,y)$ мають місце граничні співвідношення:

а) $\lim\limits_{x\to-\infty} F(x,y) = \lim\limits_{x\to-\infty} F(x,y) = \lim\limits_{\substack{x\to-\infty \\ y\to-\infty}} F(x,y) = 0$

(ймовірності неможливих подій);

б) $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(x,y) = 1$ (ймовірність достовірної події).

4. Знаходження одновимірних характеристик за двовимірними:

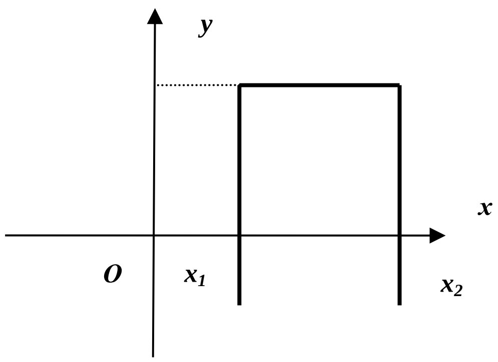
а) $\lim\limits_{y\to+\infty} F(x,y) = F_1(x)$ - функція розподілу компоненти *X;*

б) $\lim\limits_{x\to+\infty} F(x,y) = F_2(x)$ - функція розподілу компоненти Y.

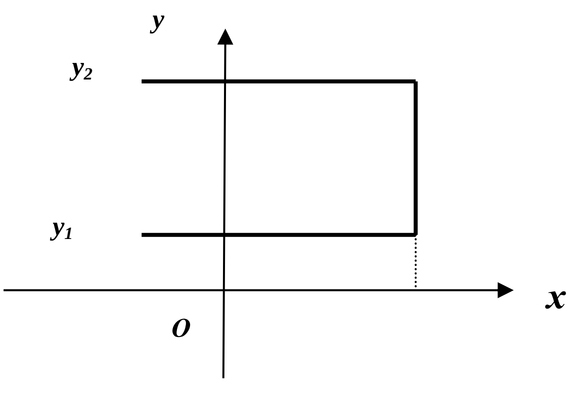
5. Ймовірність попадання випадкової точки в напівсмугу знаходиться за формулами (мал. 16 а, 16 б);

а) $P(x_1 < X < x_2, Y < y) = F(x_2, y) - F(x_1, y)$;

б) $P(X < x, y_1 < Y < y_2) = F(x, y_2) - F(x, y_1)$.



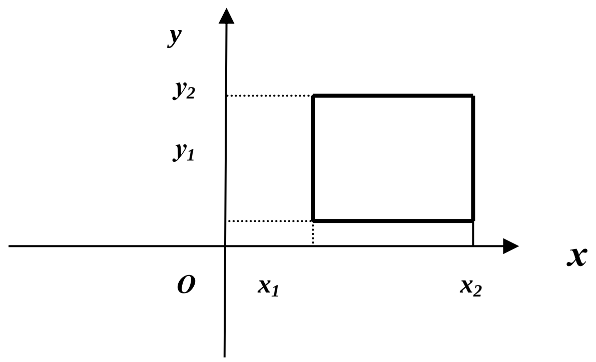
Мал. 16 а



Мал. 16 б

6. Ймовірність попадання випадкової точки *(X,Y)* в прямокутник: $(x_1 \le X \le x_2; y_1 \le Y \le y_2)$ знаходиться за формулою (див. мал. 17):

$$P(x_1 \le X \le x_2; y_1 \le Y \le y_2) = (F(x_2,y_2) - F(x_1,y_2)) - (F(x_2,y_1) - F(x_1,y_1)) \qquad (88)$$



Мал. 17

Приклад 87 . Знайти ймовірність попадання випадкової точки (X, Y) в прямокутник, обмежений лініями: $x = \pi/6, x = \pi/2, y = \pi/4, y = \pi/6,$ якщо функція розподілу $F(x, y) = sin\, x\, sin\, y\ \ (0 \le x \le \frac{\pi}{2}, 0 \le y \le \frac{\pi}{2})$.

Розв'язування. Використовуючи формулу (88), знаходимо:

$$P\left(\frac{\pi}{6} < X < \frac{\pi}{2}; \frac{\pi}{4} < Y < \frac{\pi}{3}\right) = \left(sin\frac{\pi}{2} sin\frac{\pi}{3} - sin\frac{\pi}{6} sin\frac{\pi}{3}\right) - \left(sin\frac{\pi}{2} sin\frac{\pi}{4} - sin\frac{\pi}{6} sin\frac{\pi}{4}\right) =$$

$$= \left(\frac{\sqrt{3}}{2} - \frac{1}{2}\frac{\sqrt{3}}{2}\right) - \left(\frac{\sqrt{2}}{2} - \frac{1}{2}\frac{\sqrt{2}}{2}\right) \approx 0{,}08.$$

Приклад 88. Випадковий вектор $(X; Y)$ має функцію розподілу:

$$F(x, y) = \left(\frac{1}{2} + \frac{1}{\pi} actgx\right)\left(\frac{1}{2} + \frac{1}{\pi} actgy\right).$$

Необхідно: а) обчислити $P\{-1 \le X < 1, -1 \le Y < 1\}$;

б) знайти функцію розподілу кожної компоненти вектора;

Розв'язування. а) згідно формули (88), отримаємо:

$$P\{-1 \le X < 1, -1 \le Y < 1\} = \left(\frac{1}{2} + \frac{1}{\pi} actg1\right)\left(\frac{1}{2} + \frac{1}{\pi} actg1\right) -$$

$$- \left(\frac{1}{2} + \frac{1}{\pi} actg1\right)\left(\frac{1}{2} + \frac{1}{\pi} actg(-1)\right) -$$

$$- \left(\frac{1}{2} + \frac{1}{\pi} arctg(-1)\right)\left(\frac{1}{2} + \frac{1}{\pi} arctg1\right) + \left(\frac{1}{2} + \frac{1}{\pi} actg(-1)\right)\left(\frac{1}{2} + \frac{1}{\pi} actg(-1)\right) =$$

$$= \left(\frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{4}\right)\left(\frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{4}\right) - \left(\frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{4}\right)\left(\frac{1}{2} - \frac{1}{\pi}\cdot\frac{\pi}{4}\right) - \left(\frac{1}{2} - \frac{1}{\pi}\cdot\frac{\pi}{4}\right)\left(\frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{4}\right) +$$

$$+\left(\frac{1}{2}-\frac{1}{\pi}\cdot\frac{\pi}{4}\right)\left(\frac{1}{2}-\frac{1}{\pi}\cdot\frac{\pi}{4}\right)=\frac{3}{4}\cdot\frac{3}{4}-\frac{3}{4}\cdot\frac{1}{4}-\frac{3}{4}\cdot\frac{1}{4}+\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{4}.$$

Оскільки,

$$\lim_{y\to+\infty} arctgy=\frac{\pi}{2},$$

то

$$F_X(x)=F(x,+\infty)=\left(\frac{1}{2}+\frac{1}{\pi}actgx\right)\left(\frac{1}{2}+\frac{1}{\pi}\cdot\frac{\pi}{2}\right)=\frac{1}{2}+\frac{1}{\pi}actgx.$$

Аналогічно отримаємо $F_Y(y)=\frac{1}{2}+\frac{1}{\pi}actgy$.

Щільність розподілу двовимірної випадкової величини та її властивості.

Означення. ***Щільністю розподілу ймовірностей неперервної двовимірної випадкової величини (X,Y) називається друга мішана частинна похідна від функції розподілу*** $F(x,y)$**:**

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}. \qquad (89)$$

Основні властивості щільності розподілу

1. $f(x,y)\ge 0$;
2. $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(\xi,\eta)d\xi d\eta$;
3. а) $f_1(x)=\int_{-\infty}^{+\infty} f(x,y)dy$ щільність розподілу компоненти *X*;

   б) $f_2(y)=\int_{-\infty}^{+\infty} f(x,y)dx$ щільність розподілу компоненти Y;
4. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy=1$;
5. *Ймовірність попадання двовимірної випадкової точки в довільну область D знаходиться за формулою*

$$P((X,Y)\in D)=\iint_D f(x,y)dxdy. \qquad (90)$$

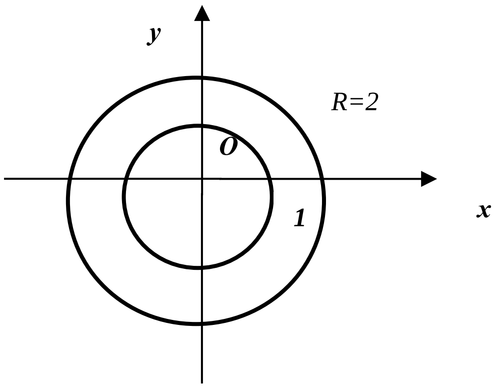
Приклад 89. У крузі $x^2+y^2\le R^2$ двовимірна щільність ймовірності задається формулою $f(x,y)=C(R-\sqrt{x^2+y^2})$, а зовні круга $f(x,y)$ = 0. Знайти: а) сталу С; б) ймовірність попадання випадкової точки *(X, Y)* у круг з радіусом r = 1 і центром у початку координат, якщо $R = 2$

Розв'язування. а) За властивістю 4 двовимірної щільності ймовірності запишемо:

$$\iint_D C\left(R-\sqrt{x^2+y^2}\right)dxdy = 1.$$

Звідси знаходимо сталу C:

$$C = \frac{1}{\iint_D (R-\sqrt{x^2+y^2})dxdy}.$$



Мал. 18

Застосувавши полярні координати $x = \rho\cos\varphi, y = \rho\sin\varphi$, обчислюємо:

$$C = \frac{1}{\int_0^{2\pi} d\varphi \int_0^R (R-\rho)\rho d\rho} = \frac{1}{2\pi\cdot\left(R\frac{\rho^2}{2}-\frac{\rho^3}{3}\right)\Big|_0^R} = \frac{3}{\pi R^3}.$$

б) За умовою $R = 2$, тому $C = {}^{3\pi}\!/_{8}$, а щільність

$$f(x,y) = \frac{3}{8\pi\left(2-\sqrt{x^2+y^2}\right)}.$$

Ймовірність попадання випадкової величини (X, Y) в круг з радіусом r =1 і центром у початку координат знаходимо за властивістю 5**.** (див. мал. 18):

$$P((X,Y)) = \frac{3}{8\pi}\iint_{D_1}\left(2-\sqrt{x^2+y^2}\right)dxdy,$$

або у полярних координатах:

$$P(X,Y) = \frac{3}{8\pi}\int_0^{2\pi} d\varphi\int_0^1 (2-\rho)\rho d\rho = \frac{3}{8\pi}\cdot 2\pi\left(2\frac{\rho^2}{2}-\frac{\rho^3}{3}\right)\Big|_0^1 = \frac{3}{4}\left(1-\frac{1}{3}\right) = \frac{1}{2}.$$

Залежні та незалежні двовимірні випадкові величини. Умовні розподіли

Випадкові величини називають незалежними, якщо закон розподілу однієї з них не залежить від того, які можливі значення набуває інша. Це означає, що ***для незалежних випадкових величини X та Y їх двовимірна функція розподілу F(X,Y) дорівнює добутку функцій розподілу компонент:***

$$F(x,y) = F_1(x) \cdot F_2(y). \tag{91}$$

Для неперервних випадкових величини X та Y їх незалежність означає, що функція щільності розподілу f(x,y) дорівнює добутку функцій щільностей компонент:

$$f(x,y) = f_1(x) \cdot f_2(y). \tag{92}$$

Розглянемо тепер двовимірну випадкову величину (X,Y), компоненти X та Y якої є залежними.

Нехай двовимірна випадкова величина (X, Y) є дискретною з такими значеннями компонент: $x_1, x_2, ..., x_n; y_1, y_2, ..., y_m$..

Позначимо умовну ймовірність того, що компонента X набуде значення x_i за умови, що $Y = y_j$, через $P\left(x_i / y_j\right), i = \overline{1,n}, j = \overline{1,m}$..

Означення. ***Для двовимірної дискретної випадкової величини (X,Y) умовним розподілом ймовірностей компоненти X при умові, що подія*** $Y = y_j$ ***настала, називається множина умовних ймовірностей*** $P\left(x_i / y_j\right), i = \overline{1,n}$, ***де***

$$P\left(x_i / y_j\right) = P(x_i, y_j) / P(y_j),\ i = \overline{1,n}.$$

Таким же чином визначається умовний розподіл ймовірностей компоненти ***Y.***

Має місце наступне твердження. ***Сума ймовірностей умовного розподілу дорівнює одиниці.***

Дійсно,

$$\sum_{i=1}^{n} P(x_i / y_j) = \sum_{i=1}^{n} P(x_i, y_j) / P(y_j) = \frac{1}{P(y_j)} \sum_{i=1}^{n} P(x_i, y_j) = \frac{P(y_j)}{P(y_j)} = 1,$$

де використано:

$$\sum_{i=1}^{n} P(x_i, y_j) = P(y_j).$$

Подібним чином встановлюється наступна формула при фіксованому x_i:

$$\sum_{j=1}^{m} P\left(y_j / x_i\right) = 1.$$

Приклад 90. Двовимірна дискретна випадкова величина (X,Y) задана таблицею розподілу

Y	X			
	1	2	3	$\sum$
4	0,10	0,15	0,20	0,45
5	0,15	0,25	0,15	0,55
$\sum$	0,25	0,40	0,35	1

Знайти:

1) умовний закон розподілу випадкової величини X при Y = 4;
2) умовний закон розподілу випадкової величини Y при X = 2.

Розв'язування. 1) Знаходимо умовний закон розподілу випадкової величини X при Y = 4:

$$p(x_1/y_1)=\frac{p_{11}}{p\{Y=y_1\}}=\frac{0{,}10}{0{,}45}=\frac{2}{9};$$

$$p(x_2/y_1)=\frac{p_{21}}{p\{Y=y_1\}}=\frac{0{,}15}{0{,}45}=\frac{3}{9};$$

$$p(x_3/y_1)=\frac{p_{31}}{p\{Y=y_1\}}=\frac{0{,}20}{0{,}45}=\frac{4}{9}.$$

Тому,

X	1	2	3
$p(X/y_1)$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{4}{9}$

2) Знаходимо умовний закон розподілу випадкової величини X при Y = 4:

$$p(y_1/x_2)\frac{p_{21}}{p\{X=x_2\}}=\frac{0{,}15}{0{,}40}=\frac{3}{8};$$

$$p(y_2/x_2)=\frac{p_{22}}{p\{X=x_2\}}=\frac{0{,}25}{0{,}40}=\frac{3}{8}.$$

Тому,

Y	4	5
$p(Y/x_2)$	$\frac{3}{8}$	$\frac{5}{8}$

Розглянемо випадок, коли (X,Y) є неперервною двовимірною випадковою величиною.

Означення. ***Умовною щільністю*** $\varphi\left(\frac{x}{y}\right)$ ***розподілу компоненти X за умови Y = y називається відношення щільності розподілу f(x,y) двовимірної випадкової величини (X,Y) до щільності розподілу*** $f_2(y)$ ***компоненти Y:***

$$\varphi(x/y) = f(x,y)/f_2(y)$$

Аналогічно визначається умовна щільність компоненти Y за умови $X = x$:

$$\psi(y/x) = f(x,y)/f_2(y).$$

Тут $\varphi(x/y) \geq 0,\ \psi(y/x) \geq 0,$ і $\int\limits_{-\infty}^{\infty}\varphi(x/y)dx = 1,\ \int\limits_{-\infty}^{\infty}\psi(x/y)dx = 1.$

Приклад 91. Випадковий вектор (X; Y) має функцію розподілу:

$$F(x,y) = \left(\frac{1}{2} + \frac{1}{\pi}actgx\right)\left(\frac{1}{2} + \frac{1}{\pi}actgy\right).$$

Необхідно:

а) визначити, залежні чи незалежні випадкові величини X и Y;

б) знайти щільність розподілу.

Розв'язування. а) оскільки $\lim\limits_{y\to+\infty} arctgy = \frac{\pi}{2}$, то отримаємо:

$$F_X(x) = F(x,+\infty) = \left(\frac{1}{2} + \frac{1}{\pi}actgx\right)\left(\frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{2}\right) = \frac{1}{2} + \frac{1}{\pi}actgx.$$

Аналогічно, $F_Y(y) = \frac{1}{2} + \frac{1}{\pi}actgy$.

Оскільки $F(x,y) = F_X(x)F_Y(y)$, то випадкові величини X и Y незалежні.

б) для знаходження щільності розподілу випадкового вектора знаходимо частинні похідні двовимірної функції розподілу:

$$\frac{\partial F}{\partial y} = \left(\frac{1}{2} + \frac{1}{\pi}actgx\right)\frac{1}{\pi}\cdot\frac{1}{1+y^2};$$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial y}\right) = \frac{1}{\pi}\cdot\frac{1}{1+x^2}\frac{1}{\pi}\cdot\frac{1}{1+y^2} = \frac{1}{\pi^2(1+x^2)(1+y^2)}.$$

§2. Числові характеристики двовимірної випадкової величини

Двовимірні випадкові величини характеризуються математичними сподіваннями та дисперсіями їх координат *M(X),M(Y),D(X),D(Y)*. Крім того, вводять характеристики спільного зв'язку координат, а саме кореляційний момент μ_{xy} і коефіцієнт кореляції r_{xy}.

Означення. ***Кореляційнім моментом двовимірної випадкової величини (X,Y) називається величина:***

$$\mu_{xy} = M\big((X - M(X))(Y - M(Y))\big)\ ,$$

або

$$\mu_{xy} = M(XY) - M(X)M(Y)\,. \tag{93}$$

Кореляційний момент має наступні властивості:

1. Якщо X і Y незалежні, то μ_{xy} =0.

   Дійсно, $M(XY)= M(X)M(Y)$, тому μ_{xy}=0.

   Зауважимо, що з рівності μ_{xy}= 0 не завжди слідує незалежність координат X і Y.

2. $\left|\mu_{xy}\right| \le \sigma(X)\sigma(Y)$, де $\sigma(X) = \sqrt{D(X)}$ та $\sigma(Y) = \sqrt{D(Y)}$ – середньоквадратичні відхилення координат X і Y.

3. При X=Y кореляційний момент μ_{xy} рівний дисперсії випадкової величини X:

$$\mu_{xx} = DX\ .$$

Означення. ***Коефіцієнт кореляції двовимірної випадкової величини (X,Y) визначається наступним чином:***

$$r_{xy} = \mu_{xy} / \sigma_x\, \sigma_y\quad . \tag{94}$$

Із означення слідує, що r_{xy} - безрозмірна величина, така, що $\left|r_{xy}\right| \le 1$.

Зауваження. Якщо $r_{xy} \ne 0$, то випадкові величини X та Y називають ***корельованими,*** а якщо r_{xy} = 0, то – ***некорельованими.*** Незалежні величини завжди некорельовані, оскільки для них виконується співвідношення: $\mu_{xy} = r_{xy} = 0$.

Математичні сподівання двовимірної випадкової величини (X,Y) задають координати центру розподілу:

$$m_x = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\, f(x,y)dxdy\,;\ m_y = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\, f(x,y)dxdy\,.$$

Дисперсії $D(X)$ та $D(Y)$ характеризують розсіювання двовимірної випадкової величини (X,Y) вздовж координатних вісей Ox та Oy і знаходяться за формулами:

$$D_x = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2\ f(x,y)dxdy - m_x^2\ ,\ \ D_y = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2\ f(x,y)dxdy - m_y^2\ .$$

Для двовимірних випадкових величин також вводять поняття умовних математичних сподівань. Наведемо відповідне означення для дискретної двовимірної випадкової величини.

Означення. ***Для дискретної двовимірної випадкової величини умовним математичним сподіванням координати Y, при умові, що координата X набула значення x, називається величина:***

$$M(Y/X=x)=\sum_{i=1}^{m} y_i\, P(y_i/x).$$

Аналогічно визначається: ***умовне математичне сподівання координати X при умові, що координата Y набула значення y:***

$$M(X/Y=y)=\sum_{i=1}^{m} x_i\, P(x_i/y).$$

Приклад 92. Двовимірна дискретна випадкова величина (X,Y) задана таблицею розподілу:

Y	X: 1	2	4	∑
0	0,1	0	0,1	0,2
2	0	0,3	0,3	0,6
5	0,2	0	0	0,2
∑	0,3	0,3	0,4	1

Знайти коефіцієнт кореляції (X,Y).

Розв'язування. Додаючи стовбці та рядки таблиці розподілу, знайдемо закони розподілу X і Y відповідно:

X	1	2	4
p	0,3	0,3	0,4

Y	0	2	5
p	0,2	0,6	0,2

Знаходимо числові характеристики координат:

$M(X) = 1\cdot 0,3 + 2\cdot 0,3 + 4\cdot 0,4 = 2,5;$

$M(X^2) = 1\cdot 0,3 + 4\cdot 0,3 + 16\cdot 0,4 = 7,9;$

$D(X) = M(X^2) - (M(X))^2 = 7,9 - 2,5^2 = =1,65;$

$\sigma(X) = \sqrt{D(X)} = 1,2845;$

$M(Y) = 0\cdot 0,2 + 2\cdot 0,6 + 5\cdot 0,2 = 2,2;$

$M(Y^2) = 0\cdot 0,2 + 4\cdot 0,6 + 25\cdot 0,2 = 7,4;$

$D(Y) = M(Y^2) - (M(Y))^2 = 7,4 - 2,2^2 = =2,56;$

$\sigma(Y) = \sqrt{D(Y)} = 1,6.$

Визначаємо характеристики сукупного зв'язку координат — кореляційний момент та коефіцієнт кореляції:

$M(X \cdot Y) = 4 \cdot 0{,}3 + 8 \cdot 0{,}3 + 5 \cdot 0{,}2 = 4{,}6;$

$\mu_{xy} = M(X \cdot Y) - M(X)M(Y) = 4{,}6 - 2{,}5 \cdot 2{,}2 = -0{,}9;.$

$$r(X,Y) = \frac{\mu_{xy}}{\sigma(X)\sigma(Y)} = \frac{-0{,}9}{1{,}2845 \cdot 1{,}6} = -0{,}438.$$

Оскільки коефіцієнт кореляції відмінний від нуля, то робимо висновок, що координати даного двовимірного вектора є залежними випадковими величинами.

§3. Функція випадкового аргументу

Функція випадкового аргументу та її розподіл

Означення. ***Випадкова величина Y називається функцією випадкового аргументу X якщо кожному можливому значенню випадкової величини X відповідає одне можливе значення Y, тобто*** $Y = \varphi(X)$**.**

Слід зазначити, що для незалежних випадкових величин X_1 та X_2 незалежними є також і випадкові величини $Y_1 = \phi_1(X_1), Y_2 = \phi_2(X_2)$.

Наведемо кілька прикладів функцій випадкового аргументу.

1. Нехай різним можливим значенням *X* відповідають різні можливі значення функції *Y*. Тоді ймовірності відповідних значень *X* та *Y* є однаковими: $P(x_k) = P(y_k), k = \overline{1,n}$.

Приклад 93. Задана дискретна випадкова величина X:

X	*1*	*4*	*7*
p	*0,25*	*0,35*	*0,4*

Знайдемо розподіл функції $Y = X^2$.

Розв'язування. Знаходимо шуканий розподіл:

$Y = X^2$	*1*	*16*	*49*
p	*0,25*	*0,35*	*0,4*

2. Нехай різним можливим значенням дискретної випадкової величини *X* відповідають значення *Y*, серед яких є рівні між собою. В цьому випадку слід додавати ймовірності значень випадкової величини *X*, які повторюються.

Приклад 94. Дискретна випадкова величина задана розподілом:

X	*-3*	*-1*	*0*	*1*	*3*
p	*0,20*	*0,10*	*0,25*	*0,15*	*0,30*

Знайдемо розподіл функції $Y = X^2$.

Розв'язування. Випадкова величина $Y = X^2$ приймає значення 9;1;0;1;9. Бачимо, що різним значенням *X*= -3 та *X*=3, а також значенням *X*=-1 та *X*=1 відповідають однакові значення випадкової величини *Y*. Додаючи ймовірності несумісних подій, маємо значення ймовірностей: *P(Y=9)=0,20+0,30=0,50*, *P(Y=1)=0,10+0,15=0,25*. Тому шуканий закон розподілу функції $Y = X^2$ має вигляд:

$Y = X^2$	0	1	9
p	0,25	0,25	0,50

Приклад 95. Випадкова величина *X* має закон розподілу:

X	0	1	2	3
p	0,1	0,3	0,4	0,2

Знайти закон розподілу випадкової величини $Y = \sin\left(\frac{\pi x}{2}\right) + 1$.

Розв'язування. Знаходимо значення функції $g(x) = \sin\left(\frac{\pi x}{2}\right) + 1$ при *x* = 0, 1, 2, 3 і отримуємо: 1, 2, 1, 0. Отже, можливими значеннями випадкової величини *Y* є числа 0, 1, 2. Тепер знаходимо ймовірності: $P\{Y = 0\} = P\{X = 3\} = 0{,}2$; $P\{Y = 1\} = P\{X = 0\} + P\{X = 2\} = 0{,}1 + 0{,}4 = 0{,}5$; $P\{Y = 2\} = P\{X = 1\} = 0{,}3$.

Закон розподілу випадкової величини *Y* приймає вигляд:

Y	0	1	2
p	0,2	0,5	0,3

3. Нехай випадкова величина *X* неперервна і має щільність розподілу ймовірностей *f(x)*. Тоді використовують твердження:

При диференційовній строго монотонній функції y=f(x) -- щільності розподілу ймовірностей випадкової величини X, оберненою до якої є функція $x = \psi(y)$, щільність розподілу g(y) випадкової величини Y знаходиться за формулою:

$$g(y) = f(\psi(y)) \cdot |\psi'(y)| \,. \qquad (95)$$

Приклад 96. Нехай випадкова величина *X* розподілена нормально з нульовим математичним сподіванням та дисперсією σ^2. Визначити закон розподілу випадкової величини $Y = X^3$.

Розв'язування. Щільність розподілу ймовірностей нормально розподіленої величини *X* при *a*=0, має вигляд: Функція $y = x^3$ монотонно зростаюча, має обернену функцію $x = y^{1/3}$. Тому за формулою (95) отримуємо:

$$g(y)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{y^{\frac{2}{3}}}{2\sigma^2}}\left(y^{\frac{1}{3}}\right)',$$

або

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

Приклад 97. Випадкова величина X розподілена за показниковим законом:

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

Знайти розподіл випадкової величини

$$Y=\sqrt{X}.$$

Розв'язування. Згідно визначення функції розподілу, запишемо функцію розподілу випадкової величини $Y=\sqrt{X}$:

$$F_Y(y)=P(\sqrt{X}<y)=\begin{cases}0, & y\le 0,\\ P(X<y^2), & y>0,\end{cases}=\begin{cases}0, & y\le 0,\\ 1-e^{-y^2}, & y>0.\end{cases}$$

Оскільки, щільність рівна похідній від функції розподілу, то отримуємо:

$$f_Y(y)=F_Y'(y)=\begin{cases}0, & y\le 0,\\ 2ye^{-y^2}, & y>0.\end{cases}$$

Математичне сподівання функції випадкової величини

При заданому законі розподілу випадкової величини X знайдемо математичне сподівання функції випадкової величини $Y=\varphi(X)$.

1. Якщо x – дискретна випадкова величина з можливими значеннями $x_1,x_2,...,x_n$, з ймовірностями відповідно $p_1,p_2,...,p_n$., то математичне сподівання дискретної випадкової величини $y=\varphi(x)$ визначають за формулою:

$$M(\varphi(X))=\sum_{i=1}^{n}\varphi(x_i)p_i\,. \tag{96}$$

Приклад 98. Знайти математичне сподівання функції $Y=\varphi(X)=X^2+3X$ для дискретної випадкової величини *X*, що задана законом розподілу:

X	1	2	3
p	0,2	0,4	0,4

Розв'язування. Спочатку визначимо можливі значення випадкової величини Y: $\varphi(1)=1^2+3=4, \varphi(2)=2^2+6=10, \varphi(3)=3^2+9=18.$ Тоді за рівністю (96), отримаємо: $M(X^2+3X)=4\cdot 0{,}2+10\cdot 0{,}4+18\cdot 0{,}4=0{,}8+4+7{,}2=12.$

***2.** Якщо аргумент X є неперервною випадковою величи*ною з щільністю розподілу $f(X)$, то математичне сподівання функції $\varphi(X)$ знаходимо за формулою:

$$M(\varphi(X))=\int_{-\infty}^{+\infty}\varphi(x)\cdot f(x)\cdot dx \quad . \tag{97}$$

Якщо можливі значення випадкової величини X належать інтервалу $(a;b)$, то (97) прийме вигляд:

$$M(\varphi(X))=\int_{a}^{b}\varphi(x)\cdot f(x)\cdot dx \quad . \tag{98}$$

Приклад 99. Неперервну випадкову величину X задано щільністю розподілу $f(x)$:

$$f(x)=\begin{cases}3x^2, \; x\in(0;1),\\ 0, \quad x\notin(0;1).\end{cases}$$

Знайти математичне сподівання функції $Y=\varphi(X)=X^2$.

Розв'язування. Скористаємося рівністю (98) і знайдемо:

$$M(\varphi(X))=M(X^2)=\int_0^1 x^2 3x^2 dx = 3\int_0^1 x^4 dx = \frac{3}{5}x^5\Big|_0^1=\frac{3}{5} \quad .$$

Дисперсія функції випадкової величини

При заданому законі розподілу випадкової величини X знайдемо дисперсію функції випадкової величини $Y=\varphi(X)$.

Відомо, що дисперсія знаходиться за формулою $D(Y)=M(Y-M(Y))^2$, Для обчислення дисперсії також можна користуватись рівністю $D(Y)=M(Y^2)-M^2(Y)$. Звідси отримуємо формулу для обчислення дисперсії функції випадкової величини $\varphi(X)$:

$$D(\varphi(X))=M(\varphi^2(X))-M^2(\varphi(X)). \tag{99}$$

Для дискретних випадкових величин X рівність (99) має вигляд:

$$D(\varphi(X))=\sum_{k=1}^{n}p_k\varphi^2(x_k)-\left(\sum_{k=1}^{n}p_k\varphi(x_k)\right)^2 \quad . \tag{100}$$

Якщо випадкова величина X неперервна, то

$$D(\varphi(X))=\int_{-\infty}^{+\infty}\varphi^2(x)f(x)dx-\left(\int_{-\infty}^{+\infty}\varphi(x)f(x)dx\right)^2, \tag{101}$$

де $f(x)$ – щільність розподілу випадкової величини X.

Якщо значення випадкової величини X знаходяться в інтервалі $(a;b)$, то рівність (101) набуде вигляду:

$$D(\varphi(X)) = \int_a^b \varphi^2(x) f(x)dx - \left(\int_a^b \varphi(x) f(x)dx \right)^2 . \qquad (102)$$

Функції кількох випадкових аргументів. Розподіл суми незалежних випадкових величин

Нехай випадкова величина Z є функцією двох випадкових змінних X та Y: $Z=f(X,Y)$. Розглянемо на прикладах розподіл суми двох випадкових величин у дискретному і неперервному випадках.

1. ***Дискретні випадкові величини***. Для того, щоб скласти закон розподілу функції $Z=X+Y$, потрібно знайти всі можливі значення Z та їх ймовірності.

Приклад 100. Розглянемо дискретні незалежні випадкові величини, які задані розподілами:

Y	*0*	*1*	*3*
p	*0,5*	*0,375*	*0,125*

X	*0*	*1*
p	*0,2*	*0,8*

Знайти розподіл суми випадкових величин $Z=X+Y$.

Розв'язування. Значення Z є сумами можливих значень випадкової величини X з усіма можливими значеннями випадкової величини Y. Тобто Z приймає значення: *0,1,2,3,4.* Знайдемо їх ймовірності. Для того, щоб $Z=0$, достатньо X набути значення $x_1 = 0$ і Y – значення $y_1 = 0$. Ці події $(X=0 \text{ і } Y=0)$ є незалежними, тому за теоремою добутку ймовірність їх сумісної появи дорівнює $0{,}5 \cdot 0{,}2 = 0{,}1$.

Аналогічно знаходимо: $P(Z=1+1=2) = 0{,}375 \cdot 0{,}8 = 0{,}3$;

Знайдемо ймовірність $P(Z=1)$. Оскільки значення $Z=1$ отримується при $(X=0 \text{ і } Y=1)$ та $(X=1 \text{ і } Y=0)$, а ці події є несумісні, то

$$P(Z=1) = P(X=0, Y=1) + P(X=1, Y=0) = 0{,}375 \cdot 0{,}2 + 0{,}5 \cdot 0{,}8 = 0{,}475 .$$

Отже шуканий розподіл буде:

Z	*0*	*1*	*2*	*3*	*4*
p	*0,10*	*0,475*	*0,30*	*0,025*	*0,10*

(перевірка: 0,10+0,475+0,30+0,025+0,10=1).

2. ***Нехай X та Y – неперервні випадкові величини***. Якщо X та Y – незалежні випадкові величини, то щільність розподілу $g(z)$ суми $Z=X+Y$ може бути знайдена за формулами:

$$g(z)=\int_{-\infty}^{+\infty} f_1(x)\,f_2(z-x)dx \quad ,$$

або

$$g(z)=\int_{-\infty}^{+\infty} f_1(z-y)\,f_2(y)dy \quad ,$$

тут $f_1; f_2$ - щільності відповідних розподілів.

Зауважимо, що для незалежних і розподілених нормально випадкових величин $X=N_1(a_1,\sigma_1)$ і $Y=N_2(a_2,\sigma_2)$ випадкова величина $Z=X+Y$ буде розподілена теж за нормальним законом $Z=N(a_1+a_2,\sqrt{\sigma^2{}_1+\sigma^2{}_2})$

Завдання для самостійної роботи студентів

1. Випадковий вектор розподілений за таким законом:

Y	*X*		
	-2	*0*	*3*
-1	*0,10*	*0,15*	*0,20*
2	*0,15*	*0,25*	*0,15*

Знайти:

а) закони розподілу кожної випадкової величини X та Y;

б) кореляційний момент та коефіцієнт кореляції;

в) умовний закон розподілу випадкової величини X за умови, що випадкова величина Y прийме значення $y_1=-1$;

г) умовний закон розподілу випадкової величини Y за умови, що випадкова величина X прийме значення $x_3=3$.

2. Випадковий вектор розподілений за таким законом:

Y	*X*		
	2	*5*	*8*
0,4	*0,15*	*0,30*	*0,35*
0,8	*0,05*	*0,12*	*0,03*

Знайти:

а) закони розподілу кожної з випадкових величин X та Y;

б) кореляційний момент та коефіцієнт кореляції;

в) умовний закон розподілу випадкової величини X за умови, що випадкова величина Y прийме значення $y_2=0,8$;

г) умовний закон розподілу випадкової величини Y за умови, що випадкова величина X прийме значення $x_1=2$.

3. Кожен з трьох відвідувачів магазину може зробити покупку з імовірністю p. Нехай випадкова величина X – число відвідувачів, що зробили покупку, а випадкова величина Y – число відвідувачів, що не зробили покупку.
   Потрібно знайти:
   а) закон розподілу випадкового вектора $(X;Y)$;
   б) центр розсіювання, кореляційний момент та коефіцієнт кореляції.

4. Випадкові величини X та Y є дискретними і незалежними. Закони їх розподілів задані:

X	0	10	15
p	$2/5$	$1/5$	$2/5$

Y	7	14	25
p	$2/6$	$1/6$	$3/6$

   Потрібно знайти:
   а) закон їх спільного розподілу;
   б) числові характеристики випадкових величин X та Y;
   в) кореляційний момент та коефіцієнт кореляції.

5. Нехай випадкова величина X – навмання обране число з множини чисел $\{1, 2, 3,4,5,6\}$, а випадкова величина Y – навмання обране число з тієї ж множини, яке більше за друге, або рівне йому.
   Потрібно знайти:
   а) закон розподілу випадкового вектора $(X;Y)$;
   б) центр розсіювання, кореляційний момент та коефіцієнт кореляції;
   в) визначити залежність випадкових величин X та Y;
   г) побудувати умовний закон розподілу випадкової величини X за умови, що випадкова величина Y прийме значення $y_2 = 3$.

6. Випадковий вектор $(X;Y)$ має функцію розподілу:

$$F(x,y)=\begin{cases}1-e^{-2x}-e^{-3y}+e^{-2x-3y}, & \text{якщо } x\ge 0, y\ge 0\\ 0, & \text{для всіх інших точок площини}\end{cases}.$$

   Потрібно знайти:
   а) ймовірність $P\{1<X\le 3,\ 1<Y\le 2\}$;
   б) функцію розподілу кожної координати вектора;
   в) визначити залежність випадкові величини X та Y;
   г) спільну щільність розподілу;
   д) центр розсіювання, кореляційний момент та коефіцієнт кореляції.

7. Випадковий вектор (X; Y) має щільність розподілу:

$$f(x,y)=\begin{cases}\cos x\cdot\cos y, & \text{якщо } 0<x<\dfrac{\pi}{2}, 0<y<\dfrac{\pi}{2}\\ 0, & \text{для всіх інших точок площини}\end{cases}.$$

Потрібно знайти:

а) щільність розподілу кожної координати вектора;

б) визначити залежні чи незалежні випадкові величини X та Y;

в) центр розсіювання, кореляційний момент та коефіцієнт кореляції.

8. Двовимірна випадкова величина *(X, Y)* задана законом розподілу

Y \ X	*25*	*32*	*53*	*60*
1,3	*0,05*	*0,12*	*0,08*	*0,04*
2,7	*0,09*	*0,30*	*0,11*	*0,21*

Знайти закони розподілу компонент X та Y.

9. Функція розподілу двовимірної випадкової величини задана формулою

$$F(x,y)=\begin{cases}\left(1-e^{-3x}\right)\cdot\left(1-e^{-y}\right) & \text{при } x\ge 0, y\ge 0;\\ 0 & \text{при } x<0 \text{ або } y<0.\end{cases}$$

Знайти функції розподілу $F_1(x)$ та $F_2(y)$ компонент X і Y.

10. Задано функцію розподілу двовимірної випадкової величиний: $F(x,y)=\frac{1}{2}\left[\sin x+\sin y-\sin(x+y)\right]$ у квадраті $0\le x\le\frac{\pi}{2}, 0\le y\le\frac{\pi}{2}$; зовні квадрата $F(x,y)=0$. Знайти щільність розподілу системи.

11. Задано щільність розподілу системи випадкових величин (X, Y)

$$f(x,y)=\frac{20}{\pi^2(16+x^2)(25+y^2)}.$$

Знайти функцію розподілу системи.

12. Випадкова величина має нормальний закон розподілу с параметрами $a=0$ и $\sigma^2=1$. Знайти закон розподілу випадкової величини $Y=X^2$.

13. Дискретна випадкова величина задана законом розподілу:

X	*-3*	*-2*	*-1*	*0*	*1*	*2*	*3*
p	*0,1*	*0,2*	*0,1*	*0,1*	*0,2*	*0,2*	*0,1*

Знайти закон розподілу випадкової величини:

а) $Y = 2X + 3$;

б) $Z = X^2 - 1$.

14. Дискретна випадкова величина задана законом розподілу:

X	$\pi/4$	$\pi/2$	$3\pi/4$
p	0,3	0,5	0,2

Знайти закон розподілу випадкових величин:

а) $Y = \cos X$;

б) $Z = \sin X$.

15. Задана щільність розподілу випадкової величини *X*:

$$f(x) = \frac{1}{\pi(1 + x^2)}, \quad -\infty < x < \infty .$$

Знайти щільність розподілу випадкової величини $Y = arctgX$.

16. Випадкова величина *X* має рівномірний розподіл на інтервалі $(0, \pi)$. Знайти щільність розподілу випадкової величини $Y = \cos X$.

17. Випадкова величина *X* має нормальний розподіл з параметрами (0;1). Знайти щільність розподілу випадкової величини $Y = e^X$.

18. Випадкова величина *X* має нормальний розподіл з параметрами (0;1). Знайти щільність розподілу випадкової величини $Y = |X|$.

19. Випадкові величини *X* та *Y* є дискретними і незалежними. Закони їх розподілу мають вигляд:

X	*0*	*1*	*2*
p	$3/8$	$3/8$	$2/8$

Y	*0*	*1*	*2*
p	$1/3$	$1/3$	$1/3$

Знайти закон розподілу випадкової величини *Z* = *X* + *Y*.

20. Випадкові величини *X* та *Y* є дискретними і незалежними. Закони їх розподілу мають вигляд:

X	*0*	*1*	*3*
p	$4/8$	$3/8$	$1/8$

Y	*0*	*1*
p	$1/3$	$2/3$

Знайти закон розподілу випадкової величини *Z* = *X* + *Y*.

21. Нехай випадкові величини X_1 та X_2 незалежні і мають рівномірний та показниковий розподіли:

$$f_{X_1}(x)=\begin{cases}1, & x\in[0,1];\\ 0, & x\notin[0,1],\end{cases}\qquad f_{X_2}(x)=\begin{cases}e^{-x}, & x\ge 0;\\ 0, & x<0.\end{cases}\quad .$$

Знайти, з використанням формул (93), (94), щільність розподілу суми випадкових величин $f_{X_1+X_2}(x)$.

22. Дискретна випадкова величина *X* задана законом розподілу:

X	*-2*	*-1*	*0*	*1*	*2*
p	*0,2*	*0,1*	*0,1*	*0,3*	*0,3*

Скласти закон розподілу випадкової величини $Y=2^X$.

23. Дискретна випадкова величина *X* задана законом розподілу:

X	*1*	*1/e*	*e*	e^3
p	*0,2*	*0,1*	*0,4*	*0,3*

Знайти числові характеристики випадкової величини $Y=\ln X$.

24. Дискретну випадкову величину *X* задано законом розподілу

X	*1*	*2*	*3*	*4*
p	*0,2*	*0,1*	*0,3*	*0,4*

Знайти числові характеристики функції $Y=2^X-1$.

25. Неперервну випадкову величину задано щільністю розподілу

$$f(x)=\begin{cases}0 & \text{при } x\le 1,\\ x-0{,}5 & \text{при } 1<x\le 2,\\ 0 & \text{при } x>2.\end{cases}$$

Знайти математичне сподівання та дисперсію функції $Y=\sqrt[3]{X}$.

26. Неперервну випадкову величину задано щільністю розподілу

$$f(x)=\begin{cases}0 & \text{при } x\le 0,\\ 2x & \text{при } 0<x\le 1,\\ 0 & \text{при } x>1.\end{cases}$$

Знайти математичне сподівання та дисперсію функції $Y=X^2$.

Розділ 3. Елементи математичної статистики

Глава 1. Вибірковий метод дослідження

§1. Задачі математичної статистики

Попередні розділи посібника були присвячені ймовірнісним характеристикам випадкових явищ. Для вивчення випадкових явищ у теорії ймовірностей вводяться такі поняття як: випадкові події, ймовірності випадкових подій, випадкові величини, функції розподілів та щільності розподілів випадкових величин, математичні сподівання, дисперсії та інші числові характеристики випадкових величин. При практичному аналізі та вивченні випадкових явищ ці характеристики невідомі і підлягають визначенню на підставі результатів спостережень і вимірів.

Математична статистика – це розділ математики, який вивчає методи відбору, обробки та інтерпретації результатів спостережень випадкових явищ для виявлення їх закономірностей й визначення їх ймовірнісних характеристик.

Базуючись на методах теорії ймовірності, математична статистика розв’язує задачі: оцінки невідомих ймовірностей випадкових подій; оцінки невідомих функцій розподілів випадкових величин; оцінки параметрів розподілів; виявлення і оцінки залежності випадкових величин; перевірки правдоподібності тверджень про розподіли.

Для вивчення випадкових явищ у математичній статистиці розглядаються поняття генеральної сукупності, вибіркової сукупності (вибірки), обсягу вибірки.

Під терміном ***генеральна сукупність*** будемо розуміти всю сукупність об’єктів, які вивчаються. Розрізняють дискретні та неперервні генеральні сукупності. Генеральна сукупність називається ***дискретною***, якщо значення ознаки відрізняються одне від одного не менш, ніж на деяке число, і ***неперервною***, якщо значення ознаки можуть відрізнятись на як завгодно малу величину.

Вибірковою сукупністю (вибіркою) називають випадково відібрані об’єкти генеральної сукупності, ***обсягом вибірки*** – кількість елементів вибірки. При цьому елементи вибірки повинні вірно представляти випадкове явище, тобто бути ***репрезентативними***.

Різні значення генеральної сукупності, що містяться у вибірці, називаються ***варіантами.*** Числа n_i, які показують, скільки разів з'явилась варіанта у вибірці обсягу n, називають ***абсолютними частотами варіанти***.

Відносними частотами називаються вирази:

$$\frac{n_1}{n};\ \frac{n_2}{n};\ \ldots\ \frac{n_k}{n}.$$

Послідовні суми частот

$$n_1+n_2+\ldots+n_i \quad , \qquad \frac{n_1}{n}+\frac{n_2}{n}+\ldots+\frac{n_i}{n} \quad .$$

називають ***накопиченими*** абсолютними та відносними частотами.

Мають місце співвідношення:

$$n_1+n_2+\ldots+n_k=n \quad , \qquad \frac{n_1}{n}+\frac{n_2}{n}+\ldots+\frac{n_k}{n}=1 \quad . \qquad (103)$$

Тобто, ***сума абсолютних частот рівна обсягу вибірки, а сума відносних частот рівна одиниці.***

Сукупність варіант разом з частотами їх появи називають ***варіаційним рядом.*** Використовують ***варіаційні ряди абсолютних*** та ***відносних частот***.

Наприклад, якщо вивчається деяка випадкова величина *X*, то генеральна сукупність *X* утворює всю сукупність значень випадкової величини, вибіркова сукупність (вибірка) – випадково (незалежно) отримані значення випадкової величини $(x_1, x_2, \ldots, x_n)$, n – обсяг вибірки, ***варіанти*** – різні значення випадкової величини, що містяться у вибірці.

При цьому $(x_1, x_2, \ldots, x_n)$ є незалежними, ***однаково розподіленими*** випадковими величинами.

Як приклад, наведемо результат вибіркового обстеження рівня заробітньої плати (в грн.) працівників середньої ланки певного комерційного банку:

3700	*3500*	*3640*	*3500*	*3780*	*3640*	*3700*
3750	*3700*	*3640*	*3560*	*3640*	*3500*	*3560*
3560	*3640*	*3560*	*3560*	*3700*	*3700*	*3700*
3640	*3700*	*3640*	*3750*	*3640*	*3500*	*3780*

Даний вибірковий статистичний матеріал є початковим для побудови варіаційних рядів та подальших досліджень.

§2. Варіаційні ряди

Нехай варіанти $(x_1, x_2, \ldots, x_n)$ розташовані у порядку їх зростання. Тобто проведено ***ранжування*** статистичних даних.

Таблиця, в першому рядку якої містяться різні ранжовані варіанти, а в другому – відповідні їм абсолютні частоти називається ***варіаційним рядом абсолютних частот*** (тут k – кількість ***різних*** значень елементів вибірки).

x_1	x_2	$\ldots$	x_k
n_1	n_2	$\ldots$	n_k

Таблиця, в першому рядку якої містяться варіанти, а в другому – відповідні їм відносні частоти називається ***варіаційним рядом відносних частот:***

x_1	x_2	…	x_k
$\frac{n_1}{n}$	$\frac{n_2}{n}$	…	$\frac{n_k}{n}$

Часто ці варіаційні ряди поєднують:

x_1	x_2	…	x_k
n_1	n_2	…	n_k
$\frac{n_1}{n}$	$\frac{n_2}{n}$	…	$\frac{n_k}{n}$

З побудови варіаційних рядів починають вивчення статистичних даних.

Якщо на площині в прямокутній системі координат побудувати точки $\left(x_i, \frac{n_i}{n}\right)$ (i = 1, 2, …, m) і з'єднати їх послідовно відрізками прямих, то отримаємо ламану лінію, що називається ***полігоном відносних частот.*** Відповідно по точках (x_i, n_i) будується ***полігон абсолютних частот***.

Якщо випадкова величина X неперервна, то складають ***інтервальний варіаційний ряд:*** перелік інтервалів та абсолютних (або відносних) частот варіант, що потрапили в кожний з цих інтервалів:

$[x_1, x_2)$	$[x_2, x_3)$	…	$[x_r, x_{r+1})$
n_1	n_2	…	n_k
$\frac{n_1}{n}$	$\frac{n_2}{n}$	…	$\frac{n_r}{n}$

де r – кількість інтервалів.

Графічно інтервальний розподіл зображується ***гістограмою***, яка представляє собою східчасту фігуру, що складається з прямокутників. Для гістограми відносних частот основою i-го прямокутника є i-й інтервал, висота прямокутника є такою, що його площа дорівнює $\frac{n_i}{n}$. З побудови гістограми маємо, що сумарна площа усіх прямокутників рівна 1. Для гістограми абсолютних частот висота прямокутника є такою, що площа дорівнює n_i, а сумарна площа усіх прямокутників рівна n.

На практиці використовують гістограми абсолютних і відносних частот.

За варіаційним рядом будують ***статистичну (емпіричну) функцію розподілу,*** тобто функцію $F_n^*(x) = \frac{k(x)}{n}$, де n – обсяг вибірки, а $k(x)$ – число значень у вибірці, менших за x. Статистична функція розподілу має властивості функції розподілу генеральної сукупності $F(x)$.

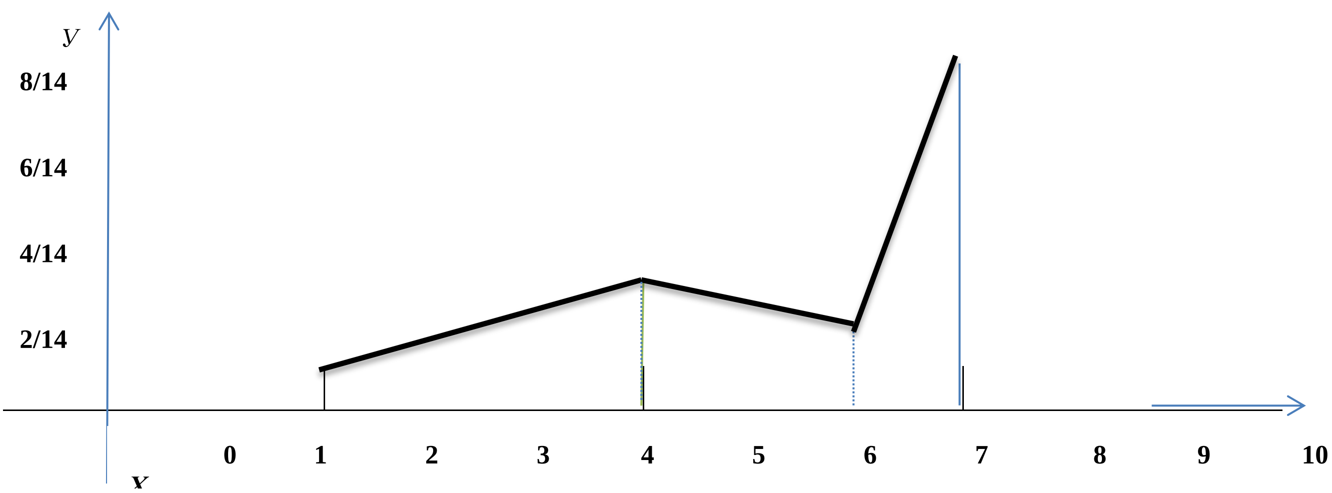
Приклад 101. Вивчається випадкова величина X – рівень відсоткових ставок по депозитах у банках міста Києва на певну дату. Отримана вибірка ставок по 14 банках: 8, 2, 8, 5, 7, 8, 8, 7, 5, 8, 8, 8, 5, 8 (%). Знайти варіаційний ряд відносних частот і статистичну (емпіричну) функцію розподілу. Побудувати полігон та емпіричну функцію розподілу.

Розв'язування. Обсяг вибірки дорівнює 14. Варіаційний ряд відносних частот має вигляд:

x_i	2	5	7	8
$\frac{n_i}{n}$	$\frac{1}{14}$	$\frac{3}{14}$	$\frac{2}{14}$	$\frac{8}{14}$

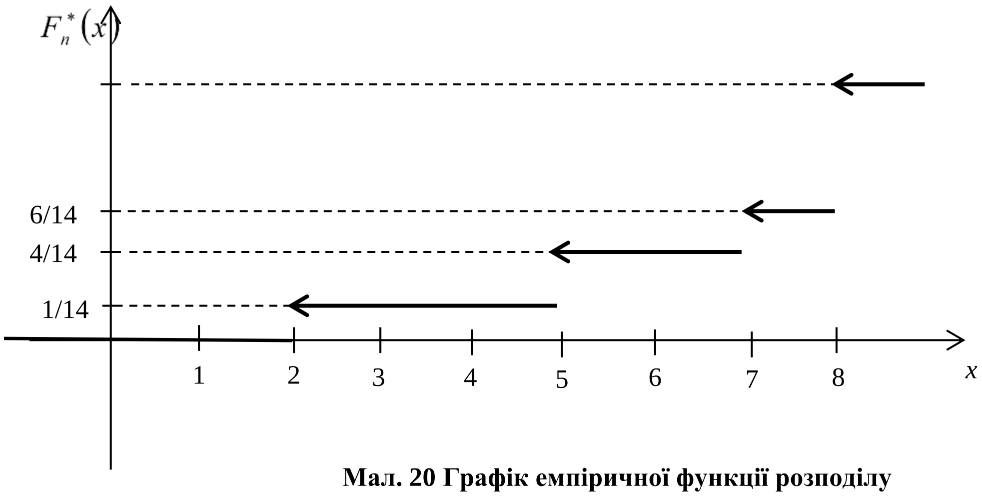
Знаходимо статистичну (емпіричну) функцію розподілу:

$$F_n^*(x) = \begin{cases} 0, & x \le 2, \\ \frac{1}{14}, & 2 < x \le 5, \\ \frac{4}{14}, & 5 < x \le 7, \\ \frac{6}{14}, & 7 < x \le 8, \\ 1, & x > 8. \end{cases}$$



Мал. 19 Графік полігону розподілу

Графік статистичної (емпіричної) функції розподілу має вигляд:



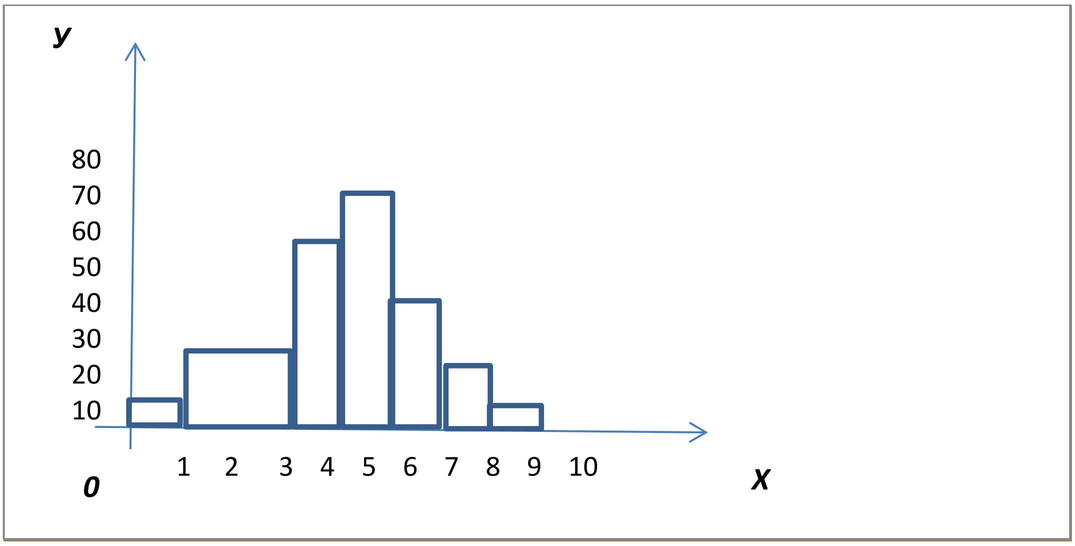
Мал. 20 Графік емпіричної функції розподілу

Приклад 102. Вивчається величина прибутку на акції металургійного підприємства. З цією метою проаналізовано дані n = 250-ти проведених торгів на біржі (результати наведено в таблиці).

інтервал	[0;1)	[1;2)	[2;3)	[3;4)	[4;5)	[5;6)	[6;7)	[7;8)
n_i	10	20	30	60	70	30	20	10

Побудувати неперервну емпіричну функцію розподілу та гістограму.

Розв`язування. Побудуємо гістограму абсолютних частот у формі східчастої фігури, що складається з прямокутників (див. мал.21). Відмітимо, що основою кожного такого прямокутника є відповідний інтервал, а висота є такою, що площа прямокутника дорівнює n_i. Побудова гістограми відносних частот проводиться так само, тільки площа кожного прямокутника дорівнює $\frac{n_i}{n}$.

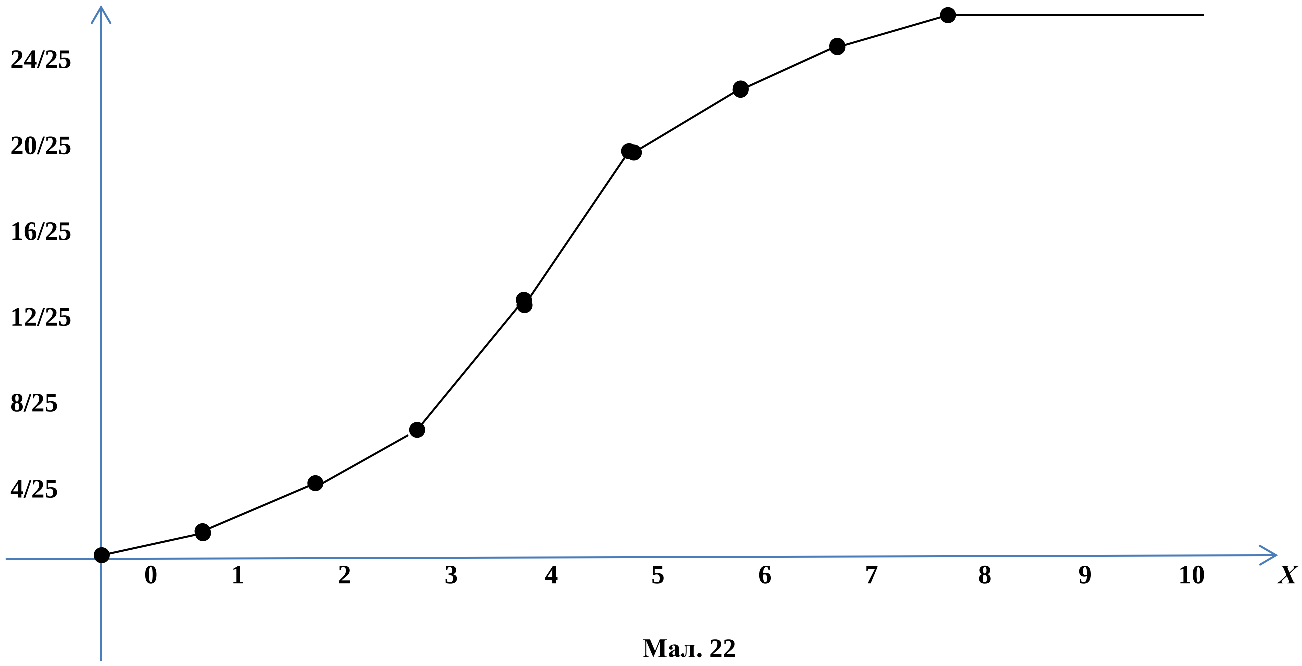


Мал. 21

Будуємо інтервальний ряд відносних частот:

інтервал	[0;1)	[1;2)	[2;3)	[3;4)	[4;5)	[5;6)	[6;7)	[7;8)
n_i	10	20	30	60	70	30	20	10
n_i/n	1/25	2/25	3/25	6/25	7/25	3/25	2/25	1/25

Далі, відкладаємо на відповідних вісях накопичені відносні частоти на інтервалах і послідовно з`єднуємо отримані точки на площині відрізками ламаних. На мал. 22 зображено графік емпіричної функції розподілу.



Мал. 22

Глава 2. Оцінки параметрів розподілів

§1. *Точкові оцінки параметрів розподілів*

Означення. ***Вибірковим середнім*** $\bar{x}$ називається середнє арифметичне значень вибірки:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ . \tag{104}$$

Означення. ***Вибірковою дисперсією*** називається величина:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{105}$$

Вибіркову дисперсію можна обчислювати за формулою аналогічною формулі обчислення дисперсії випадкової величини:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\bar{x}\right)^2 \tag{106}$$

Корінь квадратний з вибіркової дисперсії:

$$\sigma = \sqrt{S^2} \qquad (107)$$

називають ***вибірковим середнім квадратичним відхиленням***:

Якщо для X складено дискретний варіаційний ряд, то $\overline{x}$ и S^2 зручно обчислювати за формулами:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i \,,\; S^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \left(\overline{x}\right)^2. \qquad (108)$$

де n_i – абсолютні частоти варіант.

Якщо для X складено інтервальний варіаційний ряд, то $\overline{x}$ и S^2 можна обчислювати за формулами:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{r} n_i x_i^* \,,\; S^2 = \frac{1}{n}\sum_{i=1}^{r} n_i \left(x_i^*\right)^2 - \left(\overline{x}\right)^2, \qquad (109)$$

де x_i^* – середина i-того проміжку, тобто: $x_i^* = \dfrac{x_i + x_{i+1}}{2}$.

Зауважимо, що величини $F^*{}_n(x)$, $\overline{x}$, S^2 , $\sigma = \sqrt{S^2}$ є випадковими, оскільки вони залежать від випадкових значень елементів вибірки $x_1, x_2, \ldots, x_n$.

Величини $\overline{x}$, S^2, $\sigma = \sqrt{S^2}$ називають ***точковими оцінками*** невідомих параметрів $M(X), D(X), \sigma(X)$ генеральної сукупності *X*, тому що вони оцінюють ці параметри точками на числовій прямій.

Означення. ***Точковою*** оцінкою параметра a називається функція від спостережень $\tilde{a}(x_1, x_2, \ldots, x_n)$, яка знаходиться за вибіркою ($x_1, x_2, \ldots, x_n$).

Приклад 103. Для вибірки прикладу 101 знайдемо точкові оцінки параметрів $\overline{x}$, S^2 , $\sigma = \sqrt{S^2}$.

Розв'язування. Обсяг вибірки дорівнює 14. Знаходимо:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \frac{1}{14}(1\cdot 2 + 3\cdot 5 + 2\cdot 7 + 8\cdot 8) = 6{,}786 \quad (\%);$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \left(\overline{x}\right)^2 = \frac{1}{14}\left(1\cdot 2^2 + 3\cdot 5^2 + 2\cdot 7^2 + 8\cdot 8^2\right) - \left(6{,}786\right)^2 = 9{,}478;$$

$$\sigma = \sqrt{S^2} = \sqrt{9{,}478} = 3{,}079\,(\%).$$

Статистична (емпірична) функція розподілу та вибіркові характеристики є оцінками відповідно для функції розподілу та для числових характеристик генеральної сукупності (випадкової величини) *X*.

Для того, щоб точкові оцінки можна було використовувати, вони повинні задовольняти певним умовам.

Нехай a – деякий невідомий параметр розподілу.

Означення. Оцінка $\tilde{a}(x_1,x_2,...,x_n)$ параметра a називається ***незміщеною***, якщо її математичне сподівання дорівнює оцінюваному параметру:

$$M\tilde{a}(x_1,x_2,...,x_n)=a\,. \tag{110}$$

Незміщеність оцінки означає, що при заміні дійсного значення параметра його оцінкою $\tilde{a}(x_1,x_2,...,x_n)$ ми не будемо отримувати систематичних односторонніх похибок.

Означення. Оцінка $\tilde{a}(x_1,x_2,...,x_n)$ параметра a називається ***спроможною***, якщо вона збігається ***за ймовірністю*** до оцінюваного параметра:

$$\tilde{a}(x_1,x_2,...,x_n)\ \to a\ ,\quad \text{при } n\to\infty.$$

Спроможність оцінки означає, що при збільшенні обсягу вибірки оцінка наближається до оцінюваного параметра.

Зауважимо, що емпірична функція розподілу $F^*{}_n(x)$ є незміщеною і спроможною оцінкою функції розподілу $F(x)$ генеральної сукупності *X,* а вибіркове середнє $\overline{x}$ є незміщеною і спроможною оцінкою математичного сподівання $m=MX$ генеральної сукупності *X:*

$$M\overline{x}=m,\ \overline{x}\to m \tag{111}$$

Дійсно, за властивостями математичного сподівання маємо, оскільки випадкові величини $x_1,x_2,...,x_n$ однаково розподілені:

$$M\overline{x}=\frac{1}{n}M\sum_{i=1}^{n}x_i=\frac{1}{n}\sum_{i=1}^{n}Mx_i=\frac{1}{n}nm=m.$$

На відміну від вибіркового середнього $\overline{x}$, вибіркова дисперсія S^2 є ***зміщеною*** і спроможною оцінкою дисперсії $DX=\sigma^2$.

Покажемо лише ***зміщеність*** вибіркової дисперсії. Дійсно, для випадкових величин $y_i=x_i-m$ маємо:

$$\overline{y}=\overline{x}-m,\ My_i=0,\ Dy_i=My_i^2=D^2,\ MS^2=\frac{1}{n}M\sum_{i=1}^{n}(x_i-\overline{x})^2=$$

$$=\frac{1}{n}M\sum_{i=1}^{n}(y_i-\overline{y})^2=\frac{1}{n}M\sum_{i=1}^{n}y_i^2-\frac{2}{n}M(\sum_{i=1}^{n}y_i)\overline{y}+M\overline{y}^{\,2}=$$

$$=\frac{1}{n}nD^2-2M\overline{y}^{\,2}+M\overline{y}^{\,2}=D^2-M\overline{y}^{\,2}.$$

Але

$$M\overline{y}^{\,2}=M(\frac{1}{n}(y_1+y_2+...+y_n))^2=\frac{1}{n^2}\left[M\sum_{i=1}^{n}y_i^2+\sum_{i\neq j}My_iy_j\right]=\frac{1}{n^2}nD^2=\frac{D^2}{n}.$$

Тому маємо:

$$MS^2=\frac{n-1}{n}D^2. \tag{112}$$

Якщо ввести точкову оцінку – ***виправлену вибіркову дисперсію*** $\tilde{S}^2$***:***

$$\tilde{S}^2 = \frac{n}{n-1} S^2 \quad , \tag{113}$$

то співвідношення (113) означає, що виконується рівність:

$$M\tilde{S}^2 = \frac{n}{n-1} \cdot \frac{n-1}{n} D^2 = D^2. \tag{114}$$

Тобто величина $\tilde{S}^2$ буде ***незміщеною*** і ***спроможною*** оцінкою дисперсії DX. Величина $\tilde{S}^2$ називається ***виправленою вибірковою дисперсією***.

В математичній статистиці використовують і корінь квадратний з вибіркової виправленої дисперсії $\tilde{S}^2$:

$$\tilde{\sigma} = \sqrt{\tilde{S}^2} \ . \tag{115}$$

Зауважимо, що ***оцінка*** (115) ***є спроможною і зміщеною.***

§2. Інтервальне оцінювання параметрів розподілів. Довірчі інтервали та їх побудова

Точкові оцінки невідомих параметрів розподілу оцінюють ці параметри одним числом на числовій вісі і не дають можливості оцінити похибку наближення. При малих обсягах вибірок, точкове оцінювання стає ненадійним. Тому в математичній статистиці крім точкових оцінок використовують інтервальні оцінки невідомих параметрів розподілів, які дозволяють оцінювати відхилення оцінок від параметрів. Границі цих інтервалів будують на основі вибірки ($x_1, x_2, ..., x_n$), тобто вони є випадковими величинами.

Означення. ***Довірчим інтервалом з довірчою ймовірністю*** α $(0 < \alpha < 1)$ для параметра *a* генеральної сукупності *X* називається такий випадковий інтервал $(\tilde{a}_1(x_1, x_2, ..., x_n), \tilde{a}_2(x_1, x_2, ..., x_n))$, що

$$P\{\tilde{a}_1(x_1, x_2, ..., x_n) < a < \tilde{a}_2(x_1, x_2, ..., x_n)\} = \alpha \ . \tag{116}$$

Таким чином, довірчий інтервал з заданою ймовірністю *α* накриває параметр, який оцінюється.

Зауважимо, що довірчу ймовірністю *α* прагнуть обирати близькою до одиниці, базуючись на ***"принципі практичного здійснення подій, ймовірність яких є близькою до одиниці": якщо випадкова подія має ймовірність близьку до одиниці, то практично можна вважати, що в одиничному випробуванні ця подія відбудеться.*** Звичайно *α* надають значення: 0,9; 0,95; 0,99;… .

Аналогічно формулюється ***"принцип практичної неможливості малоймовірних подій": якщо випадкова подія має дуже малу ймовірність, то практично можна вважати, що в одиничному випробуванні ця подія не відбудеться.***

При цьому, як одиничне випробування розглядається експеримент по спостереженню даної вибірки. Наскільки малою повинна бути ймовірність події, щоб можна було вважати її появу в одиничному випробуванні неможливою, вирішує економіст-дослідник оскільки це залежить від конкретної задачі.

Розглянемо найбільш вживані випадки побудови довірчих інтервалів.

I. Знайдемо ***довірчий інтервал для невідомого математичного сподівання*** a ***нормально розподіленої генеральної сукупності при відомому середньому квадратичному відхиленні*** σ . Нехай $x_1, x_2, \dots, x_n$ – вибірка з нормально розподіленої генеральної сукупності. Тоді випадкова величина:

$$\frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} x_i - na\right) = \frac{1}{\sigma\sqrt{n}}\left(n\bar{x} - na\right) \quad (117)$$

розподілена нормально з параметрами (0, 1) і ймовірність попадання цієї випадкової величини в проміжок *(–t, t)* знаходимо за допомогою інтегралу Лапласа:

$$P\left\{-t < \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} x_i - na\right) < t\right\} = 2\Phi(t) \, . \quad (118)$$

Нехай довірча ймовірність $\alpha,\ 0 < \alpha < 1$ задана. Тоді рівняння (116) набуває вигляду:

$$P\left\{\left|\bar{x} - a\right| < \frac{1}{\sqrt{n}}\sigma t\right\} = 2\Phi(t) \, , \quad (119)$$

де t – розв`язок рівняння :

$$2\Phi(t) = \alpha \ . \quad (120)$$

Тобто, отримуємо довірчий інтервал для невідомого математичного сподівання a нормально розподіленої генеральної сукупності з відомим середнім квадратичним відхиленням σ у вигляді:

$$(\bar{x} - \frac{1}{\sqrt{n}}\sigma t, \bar{x} + \frac{1}{\sqrt{n}}\sigma t). \quad (121)$$

Приклад 104. При відомій дисперсії нормально розподіленої випадкової величини X рівній *0,25*; обсязі вибірки n = 25 та вибірковому середньому $\bar{x} = 52$ знайдемо для невідомого математичного сподівання a довірчий інтервал. Довірча ймовірність α рівна 0,95.

Розв’язування. Знаходимо розв’язок рівняння (120): $2\Phi(t) = 0{,}95$, використовуючи таблицю значень функції Лапласа (див. додаток 2). Для t отримуємо значення: *t=1,96*. Тепер знаходимо границі довірчого інтервалу:

$$\bar{x} - t\ \frac{\sigma}{\sqrt{n}} = 52 - 1{,}96\frac{\sqrt{0{,}25}}{\sqrt{25}} = 52 - 0{,}196 = 51{,}804,$$

$$\bar{x} + t\ \frac{\sigma}{\sqrt{n}} = 52 + 1{,}96\frac{\sqrt{0{,}25}}{\sqrt{25}} = 52 + 0{,}196 = 52{,}196.$$

Таким чином: *(51,804; 52, 196)* – шуканий довірчий інтервал з довірчою ймовірністю *0,95.*

II. Знайдемо ***довірчий інтервал для невідомого математичного сподівання a довільно розподіленої генеральної сукупності при відомому середньому квадратичному відхиленні*** σ. Нехай $x_1, x_2, \ldots, x_n$ – вибірка із генеральної сукупності з довільною функцією розподілу, невідомим середнім a і відомою дисперсією σ^2. Згідно центральної граничної теореми:

$$P\left\{-t < \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} x_i - na\right) < t\right\} \to 2\Phi(t),\ n \to \infty \quad (122)$$

Звідси отримуємо наближену рівність:

$$P\left\{\left|\bar{x} - a\right| < \frac{1}{\sqrt{n}}\sigma t\right\} \approx 2\Phi(t). \quad (123)$$

З рівності (123) знаходимо наближення до границь довірчого інтервалу вигляду (121), де t знаходиться з рівняння (120): $2\Phi(t) = \alpha$. Наближення матимуть більшу точність при збільшенні обсягу вибірки n.

III. Знайдемо ***довірчий інтервал для невідомого математичного сподівання a нормально розподіленої генеральної сукупності при невідомій дисперсії*** σ^2**.**

Доведено [1], що для генеральної сукупності – випадкової величини *N(a, σ),* яка має нормальний розподіл з параметрами *(a, σ)* випадкова величина:

$$\frac{1}{\tilde{S}}(\bar{x} - a)\sqrt{n} \quad (124)$$

розподілена за законом Стьюдента з *(n-1)*-м ступенем вільності (у формулі (124) $\tilde{S}^2$ – ***виправлена*** вибіркова дисперсія). Для розподілу Стьюдента складені таблиці значень (див. додаток 6), користуючись якими, можна для заданої довірчої ймовірності α знайти таке t_α, що виконується рівність:

$$P\left\{\left|\frac{1}{\tilde{S}}(\bar{x} - a)\sqrt{n}\right| < t_\alpha\right\} = \alpha\ . \quad (125)$$

Тоді інтервал:

$$(\bar{x} - \frac{1}{\sqrt{n}} t_\alpha \tilde{S}, \bar{x} + \frac{1}{\sqrt{n}} t_\alpha \tilde{S}) \quad (126)$$

буде ***довірчим інтервалом для невідомого математичного сподівання a з відповідною довірчою ймовірністю α.***

Зауважимо, що розподілом Стьюдента користуються при побудові довірчих інтервалів для малих обсягів вибірок.

Приклад 105. Для нормально розподіленої випадкової величини X при обсязі вибірки n = 25, ***виправленій*** дисперсії рівній 0,25 та вибірковому середньому $\bar{x} = 52$ знайдемо довірчий інтервал для невідомого математичного сподівання a. Довірча ймовірність α рівна 0,95.

Відмітимо, що даний приклад відрізняється від попереднього тим, що дисперсія генеральної сукупності невідома, а ***виправлена*** дисперсія обчислена на основі вибірки.

Розв`язування. У нашому випадку $n-1=24$, $\tilde{\sigma}=0{,}5$.

З таблиці розподілу Стьюдента (додаток 6) отримуємо значення: $t_\alpha=2{,}07$.

Тепер знаходимо границі довірчого інтервалу за формулою (126):

$$\bar{x}-t_\alpha\frac{\tilde{\sigma}}{\sqrt{n}}=52-2{,}07\frac{\sqrt{0{,}25}}{\sqrt{25}}=52-2{,}07\cdot 0{,}1=51{,}793,$$

$$\bar{x}+t_\alpha\frac{\tilde{\sigma}}{\sqrt{n}}=52+2{,}07\frac{\sqrt{0{,}25}}{\sqrt{25}}=52+0{,}207=52{,}207.$$

Таким чином: *(51,793; 52, 207)* – шуканий довірчий інтервал з довірчою ймовірністю 0,95. Відмітимо, що в даному випадку отримано ***більш широкий*** інтервал для невідомого математичного сподівання ніж у прикладі 104, де він був таким: *(51,804; 52, 196)*. Це пояснюється тим, що менше інформації було відомо про генеральну сукупність. Іншими словами, зменшення інформації про генеральну сукупність погіршило статистичні оцінки її параметрів.

Глава 3. Перевірка статистичних гіпотез

§1. Основний принцип перевірки статистичних гіпотез

Статистичною гіпотезою називається твердження про генеральну сукупність, справедливість якого перевіряється за вибіркою.

Прикладами статистичних гіпотез є: ***гіпотеза про рівність двох функцій розподілів; про вид розподілу генеральної сукупності, гіпотеза про рівність математичних сподівань двох, або кількох розподілів; про рівність дисперсій двох розподілів; про незалежність вибірок; про рівність нулю математичного сподівання генеральної сукупності*** та інші.

Справедливість статистичної гіпотези перевіряється за вибіркою. Для цього висувається ***основна гіпотеза*** H_0 та ***конкуруюча гіпотеза*** H_1. Основною гіпотезою є та гіпотеза, яка перевіряється, конкуруючою гіпотезою H_1 є гіпотеза, яка протирічить нульовій.

Якщо гіпотеза H_0 не приймається, то це означає, що приймається конкуруюча гіпотеза H_1 .

Внаслідок перевірки статистичної гіпотези можуть бути допущені помилки двох типів. ***Помилка першого типу*** полягає в тому, що буде відхилена правильна гіпотеза. Ймовірність помилки першого типу називається ***рівнем значущості*** і позначається α. ***Помилка другого типу*** полягає в тому, що буде прийнята неправильна гіпотеза. Ймовірність помилки другого типу позначається β.

Для перевірки гіпотез в математичній статистиці використовують певні випадкові величини, які називаються критеріями.

Статистичним критерієм називають випадкову величину K, яка використовується для перевірки гіпотези. Статистичний критерій будується за вибіркою $(x_1, x_2, ..., x_n)$. ***Спостереженим значенням критерия*** $\hat{K}$ називають те значення критерія, яке обчислене за вибіркою.

Ті значення вибірки $(x_1, x_2, ..., x_n)$, при яких гіпотеза приймається, називають ***областю прийняття гіпотези***; ті значення вибірки $(x_1, x_2, ..., x_n)$, при яких гіпотеза відхиляється, називають ***критичною областю***.

Критичними точками $K_{кр}$ називають точки, що відділяють критичну область від області прийняття гіпотези.

Основний принцип перевірки статистичних гіпотез полягає у тому, що гіпотезу відхиляють, якщо спостережене значення критерія належить критичній області, а якщо спостережене значення критерія належить області прийняття гіпотези, то гіпотезу приймають.

Критичну область обирають таким чином, щоб зменшити ймовірності α і β помилок двох типів. При фіксованому обсязі вибірки можна зменшувати як завгодно або лише α, або лише β. Наскільки малими повинні бути ймовірності α, або β залежить від конкретної задачі. Звичайно α надають малі значення: 0,1; 0,05; 0,01, 0,001,…

Зазначимо, що узгодження даних спостереження з гіпотезою означає лише, що ці статистичні дані не суперечать гіпотезі.

§2. Критерій узгодження χ^2. Перевірка гіпотези про закон розподілу

Нехай X – ***дискретна випадкова величина***. Перевіримо гіпотезу H_0 про те, що закон розподілу X має вигляд:

X	x_1	x_2	…	x_k
p	p_1	p_2	…	p_k

Припустимо, що варіанти $(x_1, x_2, ..., x_k)$ вибірки мають абсолютні частоти $n_1, n_2, ..., n_k$, де $\sum_{i=1}^{k} n_i = n$. Побудуємо випадкову величину χ_n^2, яка характеризує відхилення спостережуваних частот $n_1, n_2, ..., n_k$ від теоретичних ймовірностей $np_1, np_2, ..., np_k$:

$$\chi_n^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}. \tag{127}$$

Доведено [4], що при $n \to \infty$ незалежно від виду розподілу генеральної сукупності X функція χ_n^2 прямує до функції (розподілу) χ_{k-1}^2 – ***«хі-квадрат з $(k-1)$- єю ступінню вільності»***. Вважаємо, що обсяг вибірки n є досить великим, а отже можлива заміна χ_n^2 на χ_{k-1}^2.

При заданому рівні значущості α, шукаємо таке критичне значення $\chi^2_{кр} = \chi^2_\alpha$, що:

$$P\left\{\chi_{k-1}{}^2 \ge \chi^2_\alpha\right\} = \alpha . \qquad (128)$$

Тоді, значення $\chi_n{}^2$, для яких $\chi_n{}^2 \le \chi^2_\alpha$, є область допустимих значень, а значення $\chi_n{}^2$, для яких $\chi_n{}^2 > \chi^2_\alpha$, утворюють критичну область.

Зауважимо, що для граничного розподілу $\chi_{k-1}{}^2$ розв`язки рівняння (128) обчислені. Нехай обчислена величина $\tilde{\chi}_n{}^2$ – значення критерію за вибіркою.

Тоді, при

$$\tilde{\chi}_n{}^2 \le \chi^2_\alpha \qquad (129)$$

гіпотезу про вид розподілу приймаємо, як таку, що узгоджується з вибіркою, а при

$$\tilde{\chi}_n{}^2 > \chi^2_\alpha \qquad (130)$$

гіпотезу про вид розподілу відхиляємо, як таку, що не узгоджується з вибіркою.

§3. Статистична перевірка гіпотези про нормальний розподіл

Критерій узгодження χ^2 перевірки гіпотез використовується і при неперервній функції $F(x)$ розподілу генеральної сукупності. Для цього, числова вісь розбивається на інтервали, підраховуються теоретичні та спостережувані частоти для кожного інтервалу, а потім вони порівнюються за допомогою критерія χ^2.

Розглянемо у випадку неперервної генеральної сукупності X перевірку гіпотези H_0 про нормальний розподіл.

Статистична перевірка гіпотези про нормальний розподіл.

Нехай при спостереженні генеральної сукупності X отримано інтервальний варіаційний ряд:

$[x_1, x_2)$	$[x_2, x_3)$	$\dots$	$[x_r, x_{r+1})$
n_1	n_2	$\dots$	n_r

де n – обсяг вибірки, r – кількість інтервалів, n_i – абсолютні частоти попадання елементів вибірки в інтервали.

Обчислимо для цього інтервального варіаційного ряду вибіркове середнє та статистичну дисперсію за формулами (109):

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{r} n_i x_i^* , \qquad S^2 = \frac{1}{n}\sum_{i=1}^{r} n_i \left(x_i^*\right)^2 - \left(\bar{x}\right)^2 ,$$

де $x_i^* = \dfrac{x_i + x_{i+1}}{2}$ – середина i-того проміжку. Обчислимо вибіркове середнє квадратичне відхилення S.

Використовуючи критерій узгодження Пірсона, перевіримо гіпотезу H_0 про те, що неперервна генеральна сукупність X розподілена за нормальним законом з параметрами $(\bar{x}, S)$. Критерій Пірсона має вигляд:

$$\chi_n^2 = \sum_{i=1}^{r} \frac{\left(n_i - \tilde{n}_i\right)^2}{\tilde{n}_i}, \tag{131}$$

$\tilde{n}_i = n \cdot p_i$ – теоретичні абсолютні частоти, що знаходяться з припущення про нормальний з параметрами $(\bar{x}, S)$ розподіл генеральної сукупності,

$$p_i = \Phi\left(z_{i+1}\right) - \Phi\left(z_i\right), \qquad z_i = \frac{x_i - \bar{x}}{S}, \tag{132}$$

p_i – ймовірність попадання X в інтервал $\left(x_i, x_{i+1}\right)$ у припущенні, що генеральна сукупність X розподілена за нормальним законом з параметрами $(\bar{x}, S)$, $\Phi(z)$ – інтеграл Лапласа.

Доведено [4], що при $n \to \infty$, χ_n^2 (див. формулу(131)) прямує до розподілу χ_{r-3}^2 – ***«хі-квадрат з (r-3)-ма ступінями вільності»***.

Вважаємо, що обсяг вибірки n є досить великим, а отже можлива заміна χ_n^2 на χ_{r-3}^2.

Для перевірки при рівні значущості α гіпотези H_0 про нормальний розподіл генеральної сукупності X обчислимо ***спостережене значення*** критерію Пірсона за формулою:

$$\tilde{\chi}^2 = \sum_{i=1}^{r} \frac{\left(n_i - \tilde{n}_i\right)^2}{\tilde{n}_i}. \tag{133}$$

За таблицею критичних точок розподілу χ^2, по заданому рівню значущості α та кількості ступенів вільності $k = r - 3$ (r – кількість інтервалів вибірки) знаходимо критичну точку правосторонньої критичної області χ_α^2.

Якщо $\tilde{\chi}^2 < \chi_\alpha^2$, то гіпотезу H_0 про нормальний розподіл генеральної сукупності з параметрами $(\bar{x}, S)$ приймаємо, якщо $\tilde{\chi}^2 \geq \chi_\alpha^2$, то гіпотезу H_0 відхиляємо.

Зауваження. Критерій узгодження χ^2 застосовується і при довільному законі розподілу генеральної сукупності X. Доведено [4], що при $n \to \infty$, незалежно від розподілу генеральної сукупності, χ_n^2 прямує до розподілу χ_{r-s-1}^2 – ***«хі-квадрат з (r-s-1)-м ступенем вільності»***, де s – кількість параметрів теоретичного розподілу, які обчислюються за вибіркою.

Приклад 106. Використовуючи критерій згоди χ^2 , при рівні значущості α = 0,05 перевіримо гіпотезу H_0 про нормальний закон розподілу випадкової величини X, інтервальний варіаційний ряд якої наведено в таблиці (n = 500):

інтервал	*[-4;-3)*	*[-3;-2)*	*[-2;-1)*	*[-1;0)*	*[0;1)*	*[1;2)*	*[2;3)*	*[3;4)*
n_i	*10*	*30*	*70*	*130*	*120*	*90*	*40*	*10*

Розв\`язування. 1) Обчислимо вибіркове середнє $\bar{x}$ та вибіркове середнє квадратичне відхилення *S*. Проміжні обчислення запишемо в таблицю:

Середина інтервалу x_i^*	n_i	$n_i x_i^*$	$n_i\left(x_i^*\right)^2$
-3,5	*10*	*-35*	*122,5*
-2,5	*30*	*-75*	*187,5*
-1,5	*70*	*-105*	*157,5*
-0,5	*130*	*-65*	*32,5*
0,5	*120*	*60*	*30*
1,5	*90*	*135*	*202,5*
2,5	*40*	*100*	*250*
3,5	*10*	*35*	*122,5*
всього	***500***	***50***	***1105***

Знаходимо: $$\bar{x} = \frac{1}{n}\sum_{i=1}^{r} n_i x_i^* = \frac{1}{500} 50 = 0{,}1;$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{r} n_i x_i^2 - \left(\bar{x}\right)^2 = \frac{1}{500} 1105 - (0{,}1)^2 = 2{,}2; \; S = \sqrt{2{,}2} = 1{,}48.$$

2) для знаходження теоретичних частот складаємо розрахункову таблицю:

$z_i = \frac{x_i - \bar{x}}{S}$	$z_{i+1} = \frac{x_{i+1} - \bar{x}}{S}$	$\Phi(z_i)$	$\Phi(z_{i+1})$	$p_i = \Phi(z_{i+1}) - \Phi(z_i)$	np_i
$-\infty$	*-2,09*	*-0,5*	*-0,48169*	*0,01831*	*9,155*
-2,09	*-1,42*	*-0,48169*	*-0,42220*	*0,05949*	*29,745*
-1,42	*-0,74*	*-0,42220*	*-0,27035*	*0,15185*	*75,925*
-0,74	*-0,07*	*-0,27035*	*-0,02790*	*0,24245*	*121,225*
-0,07	*0,61*	*-0,02790*	*0,22907*	*0,25697*	*128,485*
0,61	*1,28*	*0,22907*	*0,39973*	*0,17066*	*85,33*
1,28	*1,96*	*0,39973*	*0,47500*	*0,07527*	*37,635*
1,96	$+\infty$	*0,47500*	*0,5*	*0,025*	*12,5*
				1,000	*500*

3) порівнюємо теоретичні та статистичні частоти. Для знаходження χ^2 складаємо розрахункову таблицю:

n_i	$n_i' = np_i$	$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$
10	*9,155*	*0,845*	*0,7140*	*0,07799*
30	*29,745*	*0,255*	*0,0650*	*0,03663*
70	*75,925*	*-5,925*	*35,1056*	*0,46237*
130	*121,225*	*8,775*	*77,0006*	*0,63519*
120	*128,485*	*-8,485*	*71,9952*	*0,00778*
90	*85,330*	*4,670*	*21,8089*	*0,25558*
40	*37,635*	*2,365*	*5,5932*	*0,14862*
10	*12,5*	*-2,5*	*6,25*	*0,5*
500	*500*	*0*		*2,12416*

Таким чином, χ^2 = 2,12416 . Кількість інтервалів r = 8, отже, число ступенів вільності дорівнює: k = r – 3 = 5. За таблицею розподілу χ^2 (див. додаток 3) знаходимо критичну точку χ^2_α=11,07. Оскільки $\chi^2 < \chi^2_\alpha$, то гіпотеза H_0 про нормальний розподіл з параметрами (0.1; 1.48) генеральної сукупності X приймається (вона не суперечить вибірковим даним).

§4. Перевірка гіпотези про істотність вибіркового коефіцієнта кореляції

Нехай вивчаються випадкові величини *X* та *Y*. Важливою є задача про з`ясування залежності, або незалежності цих випадкових величин. Будемо вважати, що спостерігається ***нормально розподілена*** двовимірна генеральна сукупність (двовимірна нормально розподілена випадкова величина) (*X, Y)* і спостережено наступні пари значень: (x_i, y_i), *i=1,2,…,n*.

x_1	x_2	…	x_n
y_1	y_2	…	y_n

З теорії ймовірностей відомо, що коефіцієнт кореляції *r* двох незалежних випадкових величин дорівнює нулю, а коефіцієнт кореляції залежних нормально розподілених випадкових величин відмінний від нуля.

Випадкові величини, для яких коефіцієнт кореляції відмінний від нуля, називаються ***корельованими,*** а випадкові величини, для яких коефіцієнт кореляції рівний нулю, називаються ***некорельованими,*** В практичних застосуваннях математичної статистики приймають некорельованість випадкових величин за ознаку їх незалежності.

Вибірковий коефіцієнт кореляції $r_в$, який будується за вибіркою, є статистичною (емпіричною) оцінкою коефіцієнта кореляції r генеральної сукупності (X,Y):

$$r_в = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\,\overline{y}}{S_X S_Y}\ , \qquad (134)$$

де

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ ,\quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad S_X{}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\ ,\ S_Y{}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2$$

– вибіркові середні та дисперсії координат X та Y.

Звичайно $r_в \neq 0$. Оскільки вибірка є випадковою, то ще не можна стверджувати, що коефіцієнт кореляції r двовимірної генеральної сукупності (X,Y) відмінний від нуля. Тобто, що X та Y є корельованими випадковими величинами.

Тому перевіряють гіпотезу H_0: ***про рівність нулю коефіцієнта кореляції нормально розподіленої генеральної сукупності:***

$$r = 0. \qquad (135)$$

Конкуруючою буде гіпотеза H_1: ***про нерівність нулю коефіцієнта кореляції:***

$$r \neq 0. \qquad (136)$$

Тобто гіпотеза про корельованість випадкових величин X та Y.

Якщо гіпотеза H_0 не приймається, то це означає, що $r_в$ істотно відхиляється від нуля. В цьому випадку кажуть, що вибірковий коефіцієнт кореляції $r_в$ є істотним, а X та Y корельовані, тобто залежні.

Якщо гіпотеза H_0 приймається, то це означає, що $r_в$ неістотний (неістотно відмінний від нуля), а X та Y некорельовані, тобто незалежні.

У якості критерія згоди приймають випадкову величину:

$$T_n = \frac{r_в\sqrt{n-2}}{\sqrt{1-r_в^2}}\ . \qquad (137)$$

Доведено [1], що якщо X та Y нормально розподілені некорельовані випадкові величини (тобто, якщо гіпотеза H_0 справджується), то T_n розподілена за ***«законом Стьюдента з (n-2)-ма ступенями вільності».***

Зауважимо, що критерій T_n може приймати від`ємні і позитивні значення. Тому критичне значення критерію T_α знаходять з вимоги, щоб ймовірність попадання значень T_n в інтервали $(-\infty, -T_\alpha) \cup (T_\alpha, +\infty)$ була не більше α . Тобто T_α знаходиться з рівняння:

$$P(|T_n| > T_\alpha) = \alpha\ . \qquad (138)$$

При знаходженні розв'язків рівняння (131) користуються таблицею «Критичні точки розподілу Стьюдента (двосторонні)» (див. додаток 5).

Після знаходження критичного значення критерію T_α знаходять значення критерію за вибіркою $\mathcal{T}_n$.

Якщо $|\mathcal{T}_n| > T_\alpha$, то гіпотеза H_0 не приймається, тобто вибірковий коефіцієнт кореляції $r_в$ є істотним, а X та Y корельовані випадкові величини.

Якщо $|\mathcal{T}_n| < T_\alpha$, то гіпотеза H_0 приймається, тобто вибірковий коефіцієнт кореляції $r_в$ є неістотним, а X та Y некорельовані випадкові величини.

Приклад 107. За вибіркою обсягу $n=62$ знайдений вибірковий коефіцієнт кореляції $r_в = 0{,}32$. З рівнем значущості $\alpha = 0{,}005$ перевірити гіпотезу про істотність вибіркового коефіцієнта кореляції $r_в$.

Розв`язування. 1) обчислюємо значення критерію за вибіркою:

$$T_{62} = \frac{r_в\sqrt{62-2}}{\sqrt{1-0{,}32^2}} = 2{,}6163.$$

2) за рівнем значущості $\alpha = 0{,}005$, користуючись таблицею «Критичні точки розподілу Стьюдента (двосторонні)» знаходимо для $(62\text{-}2=60)$-ти ступенів вільності: $T_\alpha = 2{,}00$. Оскільки виконується нерівність: $|\mathcal{T}_n| > T_\alpha$, то гіпотеза H_0 не приймається, тобто вибірковий коефіцієнт кореляції $r_в$ є істотним, а відповідно X та Y є корельованими випадковими величинами.

Глава 4. Елементи теорії кореляції та регресії

§1. Лінійна середньоквадратична регресія випадкових величин

Розглянемо дві випадкові величини (генеральні сукупності) X та Y. Ці сукупності можуть:

- бути зв'язаними певною функціональною залежністю:

$$Y=f(X), \tag{139}$$

- бути зв'язаними статистичною залежністю;
- бути незалежними.

Означення. ***Статистичною*** називається залежність, при якій зміна однієї випадкової величини веде до зміни розподілу іншої випадкової величини.

Частковим випадком статистичної залежності є кореляційна залежність.

Означення. Залежність випадкових величин є ***кореляційною***, якщо зміна однієї випадкової величини веде до зміни середнього (математичного сподівання) іншої випадкової величини.

Кореляційну залежність задають ***функціями регресії***. Функція регресії з`ясовує характер зміни в ***середньому*** випадкової величини Y – ***функції*** при можливій зміні випадкової величини X – ***аргумента.*** Тобто, X – є ***пояснююча*** змінна, а Y – є ***пояснювана*** змінна (залежна). Графік функції регресії називають ***лінією регресії***.

Найпростішим видом залежності є лінійна залежність. Тому в статистиці вводиться наступне означення.

Означення. ***Лінійною середньоквадратичною регресією «Y на X»*** називається лінійна функція:

$$y=\alpha x+\beta \,, \tag{140}$$

така, що математичне сподівання квадрата відхилення Y від $\alpha X+\beta$ приймає найменше можливе значення:

$$M(Y-\alpha X-\beta)^2 \to min. \tag{141}$$

Це означення дає спосіб знаходження невідомих параметрів (α, β) рівняння регресії (140). Тому метод знаходження рівнянь регресії шляхом мінімізації математичного сподівання квадрата відхилення отримав назву ***метода найменших квадратів (МНК).***

Таким же чином визначають і ***лінійну середньоквадратичну регресією «X на Y».***

Метод найменших квадратів застосовується також і для знаходження ***нелінійної середньоквадратичної*** регресії *«Y на X»* та *«X на Y»*. Тобто, коли функції у рівнянні (140) є нелінійними.

Зауважимо, що ***вибір (призначення) змінної, яка є аргументом і змінної, яка є функцією здійснює економіст-дослідник, спираючись на практичний зміст задачі. Математична статистика цього питання не вирішує, а лише надає досліднику засоби кількісного аналізу залежностей.***

Будемо розглядати X та Y як двовимірну генеральну сукупність (X, Y). Будемо також вважати, що ***відомі*** числові параметри $MX, MY, \sigma(X), \sigma(Y), D(X), D(Y), r_{XY}, \mu_{XY}$ генеральної сукупності (X, Y). Тобто є відомими: математичні сподівання, середні квадратичні відхилення і дисперсії координат X та Y, коефіцієнт кореляції та кореляційний момент сукупності (X, Y).

Знайдемо коефіцієнти (α, β) рівняння лінійної середньоквадратичної регресії «Y на X». Для цього перетворимо вираз $M(Y-\alpha X-\beta)^2$:

$$\begin{aligned}
&M(Y-\alpha X-\beta)^2 = M(Y-MY-\alpha(X-MX)+MY-\alpha MX-\beta)^2 = \\
&= M(Y-MY)^2+\alpha^2 M(X-MX)^2+(MY-\alpha MX-\beta)^2-2\alpha M(Y-MY)(X-MX) = \\
&= DY+\alpha^2 DX-2\alpha\, r\sqrt{DX\,DY}+(MY-aMX-b)^2 = \\
&= \left(r\sqrt{DY}-\alpha\sqrt{DX}\right)^2+\left(1-r^2\right)DY+(MY-\alpha MX-\beta)^2 .
\end{aligned} \tag{142}$$

Оскільки другий доданок не залежить від α і β, то мінімум досягається при:

$$r\sqrt{DY}-\alpha\sqrt{DX}=0,\; MY-\alpha MX-\beta=0 .$$

Звідси знаходять значення α і β:

$$\alpha = r\sqrt{\frac{DY}{DX}} = r\frac{\sigma(Y)}{\sigma(X)}, \tag{143}$$

$$\beta = MY - r\frac{\sigma(Y)}{\sigma(X)}MX .$$

Коефіцієнт α називається ***коефіцієнтом регресії «Y на X»*** .
Рівняння прямої лінії регресії «Y на X» набуває вигляду:

$$y - MY = r\frac{\sigma(Y)}{\sigma(X)}\left(x - MX\right). \tag{144}$$

Рівняння прямої лінії регресії «X на Y» набуває вигляду:

$$x - MX = r\frac{\sigma(X)}{\sigma(Y)}\left(y - MY\right). \tag{145}$$

Відповідно ***коефіцієнтом регресії «X на Y»*** називається вираз:

$$r\frac{\sigma(X)}{\sigma(Y)}. \tag{146}$$

Прямі, що визначаються на площині змінних *(x, y)* рівняннями (144), (145), проходять через точку з координатами *(MX, MY)*. Ця точка на площині називається ***центром розподілу (X, Y)***. В загальному випадку прямі регресій «*Y* на *X*» та «*X* на *Y*» не співпадають. Чим $|r|$ ближче до 1, тим більше ці прямі співпадають. Якщо $|r| \approx 0$, то прямі (144), (145) стають перпендикулярними.

Із (142) також випливає, що $\min M(Y - \alpha X - \beta)^2$ дорівнює величині:

$$\min M(Y - \alpha X - \beta)^2 = \sigma^2(Y)(1 - r^2). \tag{147}$$

Величина $\sigma^2(Y)(1 - r^2)$ дає похибку представлення *Y* через $\alpha X + \beta$. Величина $\sigma^2(Y)(1 - r^2)$ називається ***залишковою дисперсією випадкової величини Y відносно випадкової величини X.*** Якщо $|r| \approx 1$, то залишкова дисперсія близька до нуля і рівняння $y = \alpha x + \beta$ (140) майже точно представляє функціональну залежність між *Y* та *X*. Якщо $|r| \approx 0$, то вважають, що зв'язку нема. Тобто вважають, що випадкові величини *Y* та *X* є некорельованими. Якщо $|r| < 1$, то рівняння $y = \alpha x + \beta$ представляє функціональну залежність між *Y* та *X* лише наближено.

Приклад 108. За умовою прикладу 92 двовимірна дискретна випадкова величина (*X*,*Y*) задана таблицею розподілу:

Y	X			
	1	2	4	$\sum$
0	0,1	0	0,1	0,2
2	0	0,3	0,3	0,6

5	*0,2*	*0*	*0*	*0,2*
Σ	***0,3***	***0,3***	***0,4***	***1***

Скласти рівняння лінійної регресії «*Y* на *X*» та «*X* на *Y*».

Розв'язування. 1) знайдемо закони розподілу координат *X* і *Y* :

X	1	2	4
p	0,3	0,3	0,4

Y	0	2	5
p	0,2	0,6	0,2

2) Знаходимо числові характеристики координат:

$M(X) = 1{\cdot}0{,}3 + 2{\cdot}0{,}3 + 4{\cdot}0{,}4 = 2{,}5;$

$M(Y) = 0{\cdot}0{,}2 + 2{\cdot}0{,}6 + 5{\cdot}0{,}2 = 2{,}2;$

$M(X^2) = 1{\cdot}0{,}3 + 4{\cdot}0{,}3 + 16{\cdot}0{,}4 = 7{,}9;$

$M(Y^2) = 0{\cdot}0{,}2 + 4{\cdot}0{,}6 + 25{\cdot}0{,}2 = 7{,}4;$

$D(X) = M(X^2) - (M(X))^2 = 7{,}9 - 2{,}5^2 = 1{,}65;$

$D(Y) = M(Y^2) - (M(Y))^2 = 7{,}4 - 2{,}2^2 = 2{,}56;$

$\sigma(X) = \sqrt{D(X)} = 1{,}2845;$

$\sigma(Y) = \sqrt{D(Y)} = 1{,}6.$

3) Визначаємо характеристики сукупного зв'язку координат –кореляційний момент та коефіцієнт кореляції:

$$M(X{\cdot}Y) = 4{\cdot}0{,}3 + 8{\cdot}0{,}3 + 5{\cdot}0{,}2 = 4{,}6;$$

$$\mu_{XY} = M(X{\cdot}Y) - M(X)M(Y) = 4{,}6 - 2{,}5{\cdot}2{,}2 = -0{,}9;.$$

$$r_{XY} = \frac{\mu_{XY}}{\sigma(X)\sigma(Y)} = \frac{-0{,}9}{1{,}2845\cdot 1{,}6} = -0{,}438.$$

4) Знаходимо рівняння (144) регресії «*Y* на *X*»:

$$y - 2{,}2 = -0{,}438\frac{1{,}6}{1{,}2845}(x - 2{,}5) \quad \text{, або} \quad y = -0{,}5456\,x + 3{,}564.$$

5) Знаходимо рівняння (145) регресії «*X* на *Y*»:

$$x - 2{,}5 = -0{,}438\cdot\frac{1{,}2845}{1{,}6}(y - 2{,}2), \text{ або } \; x = -0{,}3516y + 3{,}2735.$$

§2. Лінійна вибіркова регресія

В попередньому параграфі були отримані для двовимірної генеральної сукупності (*X, Y)* рівняння лінійної середньоквадратичної регресії *«Y на X»* та *«X на Y»*. При отриманні цих рівнянь вважали відомими числові параметри *MX, MY,* $\sigma(X)$, $\sigma(Y)$, r_{XY} генеральної сукупності *(X, Y)*. Тобто припускали, що відомі: математичні сподівання, середні квадратичні відхилення координат *X* та *Y*, коефіцієнт кореляції та коефіцієнти регресії сукупності *(X, Y)*.

На практиці економіст–дослідник спостерігає скінчену вибірку пар значень (x_i, y_i) , $i=1,2,\ldots,n$ генеральної сукупності *(X, Y)* :

x_1	x_2	…	x_n
y_1	y_2	…	y_n

Тобто, може отримувати, спираючись на спостережені значення, лише наближення до числових параметрів та функцій регресії.

Вибірковими рівняннями регресії «Y на X» **та** ***«X на Y»*** називають рівняння регресії, в яких числові параметри генеральної сукупності *(X, Y)* оцінюються за вибіркою. Відповідні лінії на площині називають ***вибірковими лініями регресії «Y на X» та «X на Y»***.

Вибіркові рівняння регресії отримуються методом найменших квадратів.

Означення. ***Лінійною вибірковою регресією «Y на X»*** називається лінійна функція:

$$y = \hat{\alpha}\, x + \hat{\beta} \;, \tag{148}$$

така, що квадрат відхилення Y від $\hat{\alpha}\, X + \hat{\beta}$, ***обчислений за вибіркою***, приймає найменше можливе значення:

$$\sum_{i=1}^{n} (y_i - \hat{\alpha} x_i - \hat{\beta})^2 \;\to min. \tag{149}$$

Це означення дає спосіб знаходження невідомих параметрів $(\hat{\alpha}, \hat{\beta})$ вибіркового рівняння регресії (148).

Таким же чином визначають ***лінійну вибіркову регресію «X на Y»*** та ***нелінійні вибіркові регресії «X на Y»*** **та** ***«Y на X»***.

Із означення вибіркової лінії регресії випливає, що сума відстаней точок (x_i, y_i) , $i=1,2,\ldots,n$ до цієї лінії є мінімальною.

В вибіркових лінійних рівняннях регресії використовуються такі точкові оцінки *(X, Y)* як: вибіркові середні та дисперсії координат, вибіркові середні квадратичні відхилення, вибірковий коефіцієнт кореляції $r_в$:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \;,\;\; \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad S_X{}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 ,\; S_Y{}^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2 \;, \tag{150}$$

$$r_в = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}}{S_X S_Y}\,. \tag{151}$$

З умови, що коефіцієнти $(\hat{\alpha}, \hat{\beta})$ надають мінімум виразу (149), знаходимо їх значення:

$$\hat{\alpha} = r_b \frac{S_Y}{S_X}, \qquad \hat{\beta} = \bar{y} - r_b \frac{S_Y}{S_X}\bar{x} \tag{152}$$

Коефіцієнт $\hat{\alpha}$ називається ***вибірковим коефіцієнтом регресії «Y на X»***.

Вибіркове рівняння прямої лінії регресії «Y на X» набуває вигляду:

$$y - \bar{y} = r_b \frac{S_Y}{S_X}\left(x - \bar{x}\right). \tag{153}$$

Вибіркове рівняння прямої лінії регресії «X на Y» набуває вигляду:

$$x - \bar{x} = r_b \frac{S_X}{S_Y}\left(y - \bar{y}\right). \tag{154}$$

Відповідно, ***вибірковим коефіцієнтом регресії «X на Y»*** називається вираз:

$$r_b \frac{S_X}{S_Y}, \tag{155}$$

Прямі, що визначаються на площині змінних *(x, y)* рівняннями (153), (154), проходять через точку з координатами $(\bar{x}, \bar{y})$. Ця точка на площині називається ***вибірковим центром розподілу (X, Y)*** .

Чим більшим є обсяг вибірки *n,* тим точніше вибіркові рівняння і вибіркові лінії регресії представляють рівняння і лінії регресії генеральної сукупності *(X, Y).*

Приклад 109. *Компанія стільникового зв'язку досліджує залежність між витратами на рекламу нових видів послуг та прибутками, отриманими від їх впровадження. X – витрати на рекламу (в млн. грн.), Y – прибуток за певний час (в млн. грн.). Дані наведені в таблиці.*

X	*6,5*	*6,5*	*6,2*	*6,7*	*6,9*	*6,5*	*6,1*	*6,7*
Y	*105*	*125*	*110*	*120*	*140*	*135*	*95*	*130*

Необхідно:

1) знайти вибірковий коефіцієнт кореляції між *X* та *Y*;

2) при рівні значущості $\alpha = 0,05$ з`ясувати істотність вибіркового коефіцієнта кореляції;

3) скласти рівняння лінійної регресії «*Y* на *X*»;

4) спрогнозувати прибуток, який буде отримано компанією при виділенні 7,5 млн. грн. на рекламу; при виділенні 8,5 млн. грн. на рекламу.

Розв'язування. 1) Необхідні розрахунки здійснимо у таблиці:

№	x_i	y_i	$x_i\,y_i$	x_i^2	y_i^2
1.	6,5	105	682,5	42,25	11025
2.	6,5	125	812,5	42,25	15625
3.	6,2	110	682,0	38,44	12100
4.	6,7	120	804,0	44,89	14400
5.	6,9	140	966,0	47,61	19600
6.	6,5	135	877,5	42,25	18225
7.	6,1	95	579,5	37,21	9025
8.	6,7	130	871,0	44,89	16900
$\sum$	**52,1**	**960**	**6275**	**339,79**	**116900**

Звідси знаходимо:

$$\bar{x}=\frac{1}{8}\cdot 52{,}1=6{,}5125;\quad \bar{y}=\frac{1}{8}\cdot 960=120;$$

$$S_X^2=\frac{1}{8}\cdot 339{,}79-6{,}5125^2=0{,}0611;\ S_Y^2=\frac{1}{8}\cdot 116900-120^2=212{,}5;$$

$$S_X=0{,}2472;\ S_Y=14{,}5774;\ r_в=\frac{\frac{1}{8}\cdot 6275-6{,}5125\cdot 120}{0{,}2472\cdot 14{,}5774}=0{,}7978\ .$$

2) З`ясуємо істотність вибіркового коефіцієнта кореляції. Для цього обчислимо спостережене значення критерію:

$$T=\frac{r_в\sqrt{n-2}}{\sqrt{1-r_в^2}}=\frac{0{,}7978\sqrt{8-2}}{\sqrt{1-0{,}7978^2}}=7{,}94\,.$$

За таблицею критичних точок розподілу Стьюдента, за заданим рівнем значущості α = 0,05 та числом ступенів вільності r = 8–2=6 знаходимо критичну точку $T_\alpha=2{,}45$; оскільки $|T|>T_\alpha$, то гіпотеза про рівність нулю коефіцієнта кореляції відхиляється. Тобто $r_в$ є істотним і може використовуватись у побудові рівняння регресії.

4) Вибіркове рівняння прямої лінії регресії «*Y* на *X*» приймає вигляд:

$$y-120=0{,}7978\cdot\frac{14{,}5774}{0{,}2472}\cdot(x-6{,}5125),$$

звідки,

$$y=47{,}05x-186{,}39\,.$$

5) Одержане рівняння регресії можна використовувати для прогнозування можливих значень прибутку компанії щодо впровадження реклами нових видів послуг. Отримуємо прогнозні значення можливих (середніх) прибутків компанії $\overline{Y}(7{,}5);\quad \overline{Y}(8{,}5)$:

а) $\overline{Y}(7{,}5) = y(7{,}5) = 47{,}057{,}5 \cdot 7{,}5 - 186{,}39 = 166{,}485 \approx 166$ (млн. грн.);

в) $\overline{Y}(8{,}5) = y(8{,}5) = 47{,}057{,}5 \cdot 8{,}5 - 186{,}39 = 213{,}535 \approx 213$ (млн. грн.).

Завдання для самостійної роботи студентів

1. Для сукупності 6 промислових підприємств отримані дані, які характеризують залежність між вартістю основних виробничих фондів в млн. грн.: (*X*) та валовою продукцією (*Y*), також у млн. грн.: *(2,1), (4,5), (1,2), (3,3), (5,6), (8,7).* Використовуючи метод найменших квадратів визначити параметри лінійної функції, яка відображає зв'язок цих змінних.

2. Досліджується показник річного товарообігу (*Y*), та суми річних витрат (*X*) в *14*-ти філіалів торгової мережі магазинів *«АТБ»:*

Y тис.грн	*4800*	*4880*	*4720*	*4850*	*4910*	*4720*	*4800*	*4850*	*4800*	*4000*	*4880*	*4800*	*4850*	*4800*
X тис.грн	*485*	*485*	*488*	*400*	*480*	*491*	*472*	*480*	*491*	*488*	*472*	*400*	*472*	*480*

Знайти рівняння лінійної регресії «*Y* на *X*» та спрогнозувати можливі прибутки при річних витратах 600 тис. грн.

3. Знайти довірчий інтервал невідомого математичного сподівання нормально розподіленої випадкової величини *X,* якщо її середньоквадратичне відхилення $\sigma = 3$, вибіркове середнє $\overline{x} = 4$ і обсяг вибірки $n = 36$. Довірча ймовірність рівна $\alpha = 0{,}95$

4. Знайти довірчий інтервал для невідомого математичного сподівання нормально розподіленої випадкової величини *X*, якщо її середньоквадратичне відхилення $\sigma = 2{,}5$, вибіркове середнє $\overline{x} = 3{,}1$ і обсяг вибірки $n = 49$. Довірча ймовірність рівна $\alpha = 0{,}99$.

5. Побудувати полігон розподілу за заданою вибіркою:

x_i	*4*	*11*	*16*	*24*	*26*
n_i	*2*	*6*	*11*	*3*	*1*

Знайти: а) статистичну функцію розподілу та побудувати її графік;
б) вибіркове середнє та вибіркову дисперсію.

6. Побудувати гістограму за заданим розподілом вибірки:

Розмір прибутку на акцію у %	*3,50-3,60*	*3,50-3,60*	*3,50-3,60*	*3,50-3,60*
Кількість компаній	*11*	*22*	*16*	*10*

7.Знайти емпіричну функцію розподілу за заданою вибіркою:

x_i	-1	2	3	6	9
n_i	1	12	25	20	6

Обчислити вибіркове середнє та вибіркову дисперсію.

8. Знайти вибіркове середнє та вибіркову дисперсію за заданою вибіркою:

x_i	-2	-1	3	4	5
n_i	3	5	8	7	5

Побудувати:
а) статистичний розподіл частот і відносних частот;
б) полігон відносних частот;
в) статистичну функцію розподілу та її графік;

9. Досліджується залежність між собівартістю (*X*) та кількістю виготовлених виробів (*Y*) на деревообробних підприємствах:

Y тис.шт	*450*	*475*	*466*	*455*	*470*	*475*	*470*	*455*	*450*	*450*
X тис.грн	*4,5*	*4,6*	*4,7*	*4,4*	*4,5*	*4,6*	*4,7*	*4,8*	*4,3*	*4,9*

Знайти рівняння лінійної регресії «*Y* на *X»* та спрогнозувати кількість виготовлених виробів при собівартості 5300 грн.

10. За результатами вибіркового обстеження рівня вартості квитків на літаки в південному напрямку у грн. трьох авіакомпаній *A, B, C*:

A	*470*	*450*	*475*	*466*	*455*	*470*	*475*	*470*	*455*	*450*
	466	*470*	*466*	*470*	*481*	*455*	*466*	*470*	*481*	*466*
B	*592*	*588*	*580*	*584*	*575*	*584*	*592*	*584*	*588*	*592*
	588	*584*	*588*	*592*	*588*	*580*	*588*	*592*	*584*	*588*
C	*520*	*515*	*495*	*500*	*540*	*535*	*505*	*515*	*520*	*500*
	510	*500*	*505*	*490*	*525*	*515*	*490*	*525*	*510*	*505*

Побудувати: а) статистичний розподіл частот і відносних частот;
б) полігон відносних частот;
в) статистичну функцію розподілу та її графік;
г) знайти вибіркове середнє та вибіркову дисперсію.

11. Знайти довірчий інтервал для математичного сподівання *a* нормального розподілу з довірчою ймовірністю 0,95, знаючи вибіркову середню $\bar{x}$, обсяг вибірки *n* і середнє квадратичне відхилення σ:

1). $\bar{x}$ = 85,10, *n* = 169, σ = 13.
2). $\bar{x}$ = 95,04, *n* = 361, σ = 19.
3). $\bar{x}$ = 64,08, *n* = 625, σ = 15.

12. Вивчається величина прибутку акцій у вугільній промисловості. З цією метою проаналізовано дані *n* = 500 навмання відібраних акціонерів (результати наведено в таблиці). Користуючись критерієм згоди χ^2 при рівні значущості α = *0,05* перевірити гіпотезу про нормальний закон розподілу випадкової величини *X* – прибутку на акцію (для кожного з п'яти варіантів).

№ варіанту	інтервали	[0, 1)	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)	[7, 8)
1.	n_i	5	25	68	134	121	91	45	8
2.	n_i	4	26	69	135	122	92	43	7
3.	n_i	7	25	68	134	123	90	47	3
4.	n_i	8	24	71	135	124	89	45	8
5.	n_i	2	24	71	137	122	91	45	9

13. Менеджером торгівельної мережі одержано у п'ятьох філіях залежність між тривалістю *X* реалізації партії продукції (у днях) і величиною партії (у тис. шт.). Результати дослідження наведені в таблиці. Потрібно для кожної з п'яти філій:

а) знайти вибірковий коефіцієнт кореляції між *X* і *Y*;
б) знайти рівняння лінійної регресії «*Y* на *X*»;
в) перевірити гіпотезу про рівність нулю коефіцієнта кореляції при рівні значущості α = 0,01;
г) зпрогнозувати величину партії, яку можна буде реалізувати у кожній філії за 36 днів.

Філія 1.	x_i	9	12	13	15	16	19	21	22
	y_i	0,95	1,35	1,6	2	2,2	2,4	2,7	3,1
Філія 2.	x_i	11	14	15	17	18	21	23	24
	y_i	1,05	1,43	1,55	1,85	2,16	2,44	2,94	2,98
Філія 3.	x_i	13	19	21	19	20	23	25	26
	y_i	0,5	1,4	1,6	2,2	2,7	3,6	4,1	4,4
Філія 4.	x_i	16	19	20	22	23	26	28	29
	y_i	1,55	1,35	1,65	2,45	2,6	3,63	4,07	4,61
Філія 5.	x_i	22	25	26	28	29	32	34	35
	y_i	1,1	1,9	2,4	2,8	3,5	4,1	4,6	4,9

14. Побудувати рівняння регресії залежності заробітної плати (*Y)* працівників медичної галузі від віку (*X*):

Вік (роки)	*40*	*42*	*44*	*46*	*48*	*50*	*52*	*54*	*56*
Заробітна плата (грн.)	*1248*	*1287*	*1354*	*1446*	*1458*	*1544*	*1567*	*1611*	*1695*

За одержаним рівнянням регресії зпрогнозувати рівень середньої заробітної плати медпрацівника, який досягнув віку 60 років.

15. Побудувати рівняння регресії залежності врожайності зернової культури (*Y)* від кількість внесених мінеральних добрив (*X):*

Мін.добрива (кг)	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*
Врожайність, ц/ га	*29,4*	*30,2*	*32,8*	*33,4*	*33,9*	*34,3*	*34,6*	*35,7*	*36,7*	*37,7*

На основі одержаного рівняння регресіі зпрогнозувати врожайність зернової культури при внесенні 28,5 кг мінеральних добрив.

16. Підсумкове розрахункове завдання по темі «Первинний статистичний аналіз фінансових даних у банківській діяльності».

a) Створити вибірки для дискретних і неперервних ознак базуючись на реальних статистичних даних фінансової діяльності відділення банку (*статистика кредитної і депозитної діяльності, статистика кредитоспроможності клієнтів, статистичний аналіз прогнозування ринку банківських послуг та інші*).

b) Побудувати варіаційний ряд вибірки.

c) Побудувати гістограму (полігон відносних частот) та графік емпіричної функції розподілу.

d) Знайти числові характеристики вибірки: вибіркові середні, вибіркові дисперсії, вибіркові середні квадратичні відхилення.

e) Використовуючи критерій Пірсона при рівні значущості $\alpha = 0{,}05$ перевірити, чи узгоджується гіпотеза про нормальний закон розподілу вибірки.

f) Побудувати довірчі інтервали для невідомих параметрів сукупності з довірчою ймовірністю (надійністю) 0,95.

g) Знайти рівняння регресії та перевірити істотність коефіцієнта кореляції досліджуваних ознак.

h) Використовуючи рівняння регресії, зпрогнозувати можливі рівні значень досліджуваних факторів.

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

	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**	**007**	**0069**	**0067**	**0065**	**0063**	**0061**
2,9	**0060**	**0058**	**0056**	**0055**	**0053**	**0051**	**0050**	**0048**	**0047**	**0046**
3,1	**0033**	**0032**	**0031**	**0030**	**0029**	**0028**	**0027**	**0026**	**0025**	**0025**
3,2	**0024**	**0023**	**0022**	**0022**	**0021**	**0020**	**0020**	**0019**	**0018**	**0018**
3,3	**0017**	**0017**	**0016**	**0016**	**0015**	**0015**	**0014**	**0014**	**0013**	**0013**
3,4	**0012**	**0012**	**0012**	**0011**	**0011**	**0010**	**0010**	**0010**	**0009**	**0009**
3,5	**0009**	**0008**	**0008**	**0008**	**0008**	**0007**	**0007**	**0007**	**0007**	**0006**
3,6	**0006**	**0006**	**0006**	**0005**	**0005**	**0005**	**0005**	**0005**	**0005**	**0004**
3,7	**0004**	**0004**	**0004**	**0004**	**0004**	**0004**	**0003**	**0003**	**0003**	**0003**
3,8	**0003**	**0003**	**0003**	**0003**	**0003**	**0002**	**0002**	**0002**	**0002**	**0002**
3,9	**0002**	**0002**	**0002**	**0002**	**0002**	**0002**	**0002**	**0002**	**0001**	**0001**

Додаток 2. Таблиця значень функції $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	**0,0000**	**0,52**	**0,1985**	**1,04**	**0,3508**	**1,56**	**0.4406**	**2,16**	**0,4846**
0,01	**0,0040**	**0,53**	**0,2019**	**1,05**	**0,3531**	**1,57**	**0,4418**	**2,18**	**0,4854**
0,02	**0,0080**	**0,54**	**0,2054**	**1,06**	**0,3554**	**1,58**	**0,4429**	**2,20**	**0,4861**
0,03	**0,0120**	**0,55**	**0,2088**	**1,07**	**0,3577**	**1,59**	**0,4441**	**2,22**	**0,4868**
0,04	**0,0160**	**0,56**	**0,2123**	**1,08**	**0,3599**	**1,60**	**0,4452**	**2,24**	**0,4875**
0,05	**0,0199**	**0.57**	**0,2157**	**1,09**	**0,3621**	**1.61**	**0,4463**	**2,26**	**0,4881**
0,06	**0,0239**	**0.58**	**0,2190**	**1,10**	**0,3643**	**1,62**	**0,4474**	**2,28**	**0,4887**
0,07	**0,0279**	**0,59**	**0,2224**	**1,11**	**0,3665**	**1,63**	**0,4484**	**2,30**	**0,4893**
0,08	**0,0319**	**0,60**	**0,2257**	**1,12**	**0,3686**	**1,64**	**0,4495**	**2,32**	**0,4898**
0,09	**0,0359**	**0,61**	**0,2291**	**1,13**	**0,3708**	**1,65**	**0,4505**	**2,34**	**0,4904**
0,10	**0,0398**	**0,62**	**0,2324**	**1,14**	**0,3729**	**1,66**	**0,4515**	**2,36**	**0,4909**
0,11	**0,0438**	**0,63**	**0,2357**	**1,15**	**0,3749**	**1,67**	**0,4525**	**2,38**	**0,4913**
0,12	**0,0478**	**0,64**	**0,2389**	**1,16**	**0,3770**	**1,68**	**0,4535**	**2,40**	**0,4918**
0,13	**0,0517**	**0,65**	**0,2422**	**1,17**	**0,3790**	**1,69**	**0,4545**	**2,42**	**0,4922**
0,14	**0,0557**	**0,66**	**0,2454**	**1,18**	**0,3810**	**1,70**	**0,4554**	**2,44**	**0,4927**
0,15	**0,0596**	**0,67**	**0,2486**	**1,19**	**0,3830**	**1,71**	**0,4564**	**2,46**	**0,4931**
0,16	**0,0636**	**0,68**	**0,2517**	**1,20**	**0,3849**	**1,72**	**0,4573**	**2,48**	**0,4934**
0,17	**0,0675**	**0,69**	**0,2549**	**1,21**	**0,3869**	**1,73**	**0,4582**	**2,50**	**0,4938**
0,18	**0,0714**	**0,70**	**0,2580**	**1,22**	**0,3883**	**1,74**	**0,4591**	**2,52**	**0,4941**
0,19	**0,0753**	**0,71**	**0,2611**	**1,23**	**0,3907**	**1,75**	**0,4599**	**2,54**	**0,4945**
0,20	**0,0793**	**0,72**	**0,2642**	**1,24**	**0,3925**	**1,76**	**0,4608**	**2,56**	**0,4948**
0,21	**0,0832**	**0,73**	**0,2673**	**1,25**	**0,3944**	**1,77**	**0,4616**	**2,58**	**0,4951**
0,22	**0,0871**	**0,74**	**0,2703**	**1,26**	**0,3962**	**1,78**	**0,4625**	**2,60**	**0,4953**
0,23	**0,0910**	**0,75**	**0,2734**	**1,27**	**0,3980**	**1,79**	**0,4633**	**2,62**	**0,4956**
0,24	**0,0948**	**0,76**	**0,2764**	**1,28**	**0,3997**	**1,80**	**0,4641**	**2,64**	**0,4959**
0,25	**0,0987**	**0,77**	**0,2794**	**1,29**	**0,4015**	**1,81**	**0,4649**	**2,66**	**0,4961**
0,26	**0,1026**	**0,78**	**0,2823**	**1,30**	**0,4032**	**1,82**	**0,4656**	**2,68**	**0,4963**
0,27	**0,1064**	**0,79**	**0,2852**	**1,31**	**0,4049**	**1,83**	**0,4664**	**2,70**	**0,4965**
0,28	**0,1103**	**0,80**	**0,2881**	**1,32**	**0,4066**	**1,84**	**0,4671**	**2,72**	**0,4967**
0,29	**0,1141**	**0,81**	**0,2910**	**1,33**	**0,4082**	**1,85**	**0,4678**	**2,74**	**0,4969**
0,30	**0,1179**	**0,82**	**0,2939**	**1,34**	**0,4099**	**1,86**	**0,4686**	**2,76**	**0,4971**
0,31	**0,1217**	**0,83**	**0,2967**	**1,35**	**0,4115**	**1,87**	**0,4693**	**2,78**	**0,4973**
0,32	**0,1255**	**0,84**	**0,2995**	**1,36**	**0,4131**	**1,88**	**0,4699**	**2,80**	**0,4974**
0,33	**0,1293**	**0,85**	**0,3023**	**1,37**	**0,4147**	**1,89**	**0,4706**	**2,82**	**0,4976**
0,34	**0,1331**	**0,86**	**0,3051**	**1,38**	**0,4162**	**1,90**	**0,4713**	**2,84**	**0,4977**
0,35	**0,1368**	**0,87**	**0,3078**	**1,39**	**0,4177**	**1,91**	**0,4719**	**2,86**	**0,4979**
0,36	**0,1406**	**0,88**	**0,3106**	**1,40**	**0,4192**	**1,92**	**0,4726**	**2,88**	**0,4980**
0,37	**0,1443**	**0,89**	**0,3133**	**1,41**	**0,4207**	**1,93**	**0,4732**	**2,90**	**0,4981**
0,38	**0,1480**	**0,90**	**0,3159**	**1,42**	**0,4222**	**1,94**	**0,4738**	**2,92**	**0,4982**
0,39	**0,1517**	**0,91**	**0,3186**	**1,43**	**0,4236**	**1,95**	**0,4744**	**2,94**	**0,4984**
0,40	**0,1554**	**0,92**	**0,3212**	**1,44**	**0,4251**	**1,96**	**0,4750**	**2,96**	**0,4985**
0,41	**0,1591**	**0,93**	**0,3238**	**1,45**	**0,4265**	**1,97**	**0,4756**	**2,98**	**0,4986**
0,42	**0,1628**	**0,94**	**0,3264**	**1,46**	**0,4279**	**1,98**	**0,4761**	**3,00**	**0,4986**
0,43	**0,1664**	**0,95**	**0,3289**	**1,47**	**0,4292**	**1,99**	**0,4767**	**3,20**	**0,4993**
0,44	**0,1700**	**0,96**	**0,3315**	**1,48**	**0,4306**	**2,00**	**0,4772**	**3,40**	**0,4996**
0,45	**0,1736**	**0,97**	**0,3340**	**1,49**	**0,4319**	**2,02**	**0,4783**	**3,60**	**0,4998**
0,46	**0,1772**	**0,98**	**0,3365**	**1,50**	**0,4332**	**2,04**	**0,4793**	**3,80**	**0,4999**
0,47	**0,1808**	**0,99**	**0,3389**	**1,51**	**0,4345**	**2,06**	**0,4803**	**4,00**	**0,4999**
0,48	**0,1844**	**1,00**	**0,3413**	**1,52**	**0,4357**	**2,08**	**0,4812**	**4,50**	**0,4999**
0,49	**0,1879**	**1,01**	**0,3438**	**1,53**	**0,4370**	**2,10**	**0,4821**	**5,00**	**0,4999**
0,50	**0,1915**	**1,02**	**0,3461**	**1,54**	**0,4382**	**2,12**	**0,4830**		
0,51	**0,1950**	**1,03**	**0,3485**	**1,55**	**0,4394**	**2,14**	**0,4838**		

Додаток 3. Критичні точки розподілу Пірсона $\chi^2_{\alpha;n}$

n	Значення α							
	0,990	0,975	0,950	0,900	0,100	0,050	0,025	0,010
1	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,64
2	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21
3	0,12	0,22	0,35	0,58	6,23	7,82	9,35	11,34
4	0,30	0,48	0,71	1,06	7,78	9,48	11,14	13,28
5	0,55	0,83	1,14	1,61	9,24	11,07	12,83	15,09
6	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48
8	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09
9	2,90	2,70	3,32	4,17	14,68	16,92	19,02	21,67
10	2,56	3,25	3,94	4,86	15,99	18,31	20,48	23,21
11	3,05	3,82	4,58	5,58	17,28	19,68	21,92	24,72
12	3,57	4,40	5,23	6,3	18,55	21,03	23,34	26,22
13	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69
14	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14
15	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58
16	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00
17	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41
18	7,02	8,23	9,39	10,86	25,99	28,87	31,53	34,81
19	7,63	8,91	10,12	11,65	27,2	30,14	32,85	36,19
20	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57
21	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93
22	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29
23	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64
24	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,92
25	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31
26	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64
27	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96
28	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,26
29	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59
30	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89
40	22,16	24,43	26,51	29,05	51,80	55,76	59,34	63,69
50	29,71	32,36	34,76	37,69	63,17	67,50	71,42	76,15
60	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38
70	45,44	48,76	51,74	55,33	85,53	90,53	95,02	100,4
80	53,54	57,15	60,39	64,28	96,58	101,9	106,6	112,3
90	61,75	65,65	69,13	73,29	107,6	113,1	118,1	124,1
100	70,06	74,22	77,93	82,36	118,5	124,3	129,6	135,8

Додаток 4. Критичні точки розподілу Стьюдента (правостороння критична область)

Кількість ступенів вільності n	Рівень значущості α				Кількість ступенів вільності n	Рівень значущості α			
	0,050	0,025	0,010	0,005		0,050	0,025	0,010	0,005
1	6,314	12,706	31,821	63,657	18	1,734	2,101	2,552	2,878
2	2,920	4,303	6,965	9,925	19	1,729	2,093	2,539	2,861
3	2,353	3,182	4,541	5,841	20	1,725	2,086	2,528	2,845
4	2,132	2,776	3,747	4,604	21	1,721	2,080	2,518	2,831
5	2,015	2,571	3,365	4,032	22	1,717	2,074	2,508	2,819
6	1,943	2,447	3,143	3,707	23	1,714	2,069	2,500	2,807
7	1,895	2,365	2,998	3,499	24	1,711	2,064	2,492	2,797
8	1,860	2,306	2,896	3,355	25	1,708	2,060	2,485	2,787
9	1,833	2,262	2,821	3,250	26	1,706	2,056	2,479	2,779
10	1,812	2,228	2,764	3,169	27	1,703	2,052	2,473	2,771
11	1,796	2,201	2,718	3,106	28	1,701	2,048	2,467	2,763
12	1,782	2,179	2,681	3,055	29	1,699	2,045	2,462	2,756
13	1,771	2,160	2,651	3,012	30	1,697	2,042	2,457	2,750
14	1,761	2,145	2,624	2,977	40	1,684	2,021	2,423	2,704
15	1,753	2,131	2,602	2,947	60	1,671	2,000	2,390	2,660
16	1,746	2,120	2,583	2,921	120	1,658	1,980	2,358	2,617
17	1,740	2,110	2,567	2,898	∞	1,645	1,960	2,326	2,576

Додаток 5. Критичні точки розподілу Стьюдента (двостороння критична область)

Кількість ступенів вільності n	Рівень значущості α				Кількість ступенів вільності n	Рівень значущості α			
	0,10	0,05	0,02	0,01		0,10	0,05	0,02	0,01
1	6,314	12,706	31,821	63,657	18	1,734	2,101	2,552	2,878
2	2,920	4,303	6,965	9,925	19	1,729	2,093	2,539	2,861
3	2,353	3,182	4,541	5,841	20	1,725	2,086	2,528	2,845
4	2,132	2,776	3,747	4,604	21	1,721	2,080	2,518	2,831
5	2,015	2,571	3,365	4,032	22	1,717	2,074	2,508	2,819
6	1,943	2,447	3,143	3,707	23	1,714	2,069	2,500	2,807
7	1,895	2,365	2,998	3,499	24	1,711	2,064	2,492	2,797
8	1,860	2,306	2,896	3,355	25	1,708	2,060	2,485	2,787
9	1,833	2,262	2,821	3,250	26	1,706	2,056	2,479	2,779
10	1,812	2,228	2,764	3,169	27	1,703	2,052	2,473	2,771
11	1,796	2,201	2,718	3,106	28	1,701	2,048	2,467	2,763
12	1,782	2,179	2,681	3,055	29	1,699	2,045	2,462	2,756
13	1,771	2,160	2,651	3,012	30	1,697	2,042	2,457	2,750
14	1,761	2,145	2,624	2,977	40	1,684	2,021	2,423	2,704
15	1,753	2,131	2,602	2,947	60	1,671	2,000	2,390	2,660
16	1,746	2,120	2,583	2,921	120	1,658	1,980	2,358	2,617
17	1,740	2,110	2,567	2,898	∞	1,645	1,960	2,326	2,576

Додаток 6. Таблиця значень $t_\alpha = t(\alpha, n)$ для розподілу Стьюдента

	Значення α		
	0,95	0,99	0,999
5	2,78	4,60	8,61
6	2,57	4,03	6,86
7	2,45	3,71	5,96
8	2,37	3,50	5,41
9	2,31	3,36	5,04
10	2,26	3,25	4,78
11	2,23	3,17	4,59
12	2,20	3,11	4,44
13	2,18	3,06	4,32
14	2,16	3,01	4,22
15	2,15	2,98	4,14
16	2,13	2,95	4,07
17	2,12	2,92	4,02
18	2,11	2,90	3,97
19	2,10	2,88	3,92
20	2,093	2,861	3,883
25	2,064	2,797	3,745
30	2,045	2,756	3,659
35	2,032	2,720	3,600
40	2,023	2,708	3,558
45	2,016	2,692	3,527
50	2,008	2,679	3,502
60	2,001	2,662	3,464
70	1,996	2,649	3,439
80	1,991	2,640	3,418
90	1,987	2,633	3,403
100	1,984	2,627	3,392
120	1,980	2,617	3,374
∞	1,960	2,576	3,291

Список рекомендованої літератури

1. Гнеденко Б.В. Курс теории вероятностей. – М.: Наука, 1988.
2. Вища математика. Спеціальні розділи . Книга 2, 2-ге вид., за ред. проф. Кулініча Г.Л. - К.: Либідь, 2003.
3. Коваленко И.Н., Филиппова А.А. Теория вероятностей и математическая статистика. – М., 1973.
4. Гмурман В. Е. Введение в теорию вероятностей и математическую статистику. - М.: Высшая школа, 1977.
5. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике. - М: Высшая школа, 1975, 335 с.
6. Кремер Н.Ш. Теория вероятностей и математическая статистика. – М., Юнити 2004.
7. Мартиненко М.А., Клименко Р.К., Лебедева І.В. Теорія ймовірностей.- К.:УДУХТ, 1999.
8. Мишура Ю.С. Методические указания к изучению теории вероятностей и математической статистики. – К., изд-во КГУ, 1984.
9. Феллер В. Введение в теорию вероятностей и ее приложения. - М.: Мир, 1983, 527 с.

Зміст