Київський національний університет імені Тараса Шевченка

М.П. Моклячук

Лекції з теорії ймовірностей та математичної статистики

Київ 2020

М.П. Моклячук. Лекції з теорії ймовірностей та математичної статистики. – К., 2020. – 179 с.

Рецензенти: доктор фіз.-мат. наук, професор Ю.В. Козаченко,
д-р фіз.-мат. наук, проф., член-кор. НАН України П.С. Кнопов

Рекомендовано до друку вченою радою механіко – математичного факультету Київського національного університету імені Тараса Шевченка (протокол № 9 від 10 лютого 2020 року)

Викладено основи теорії ймовірностей. Теоретичний матеріал допов-нено прикладами і задачами прикладного характеру, що ілюструють за-стосування теорії.

Для студентів університетів.

Зміст

Вступ

На основі спостережень і експерименту наука приходить до формулювання закономірностей, яким підпорядковано явища, які вона вивчає. Найпростіша та найбільш поширена схема закономірностей, які наука встановлює, така:

(1). При кожному здійсненні комплексу умов відбувається подія A.

Так, наприклад, якщо вода при атмосферному тиску в 760 мм нагрівається більш ніж на 100 °C (комплекс умов Ξ), то вона перетворюється на пару (подія A).

Подія, яка неминуче відбувається при кожній реалізації комплексу умов Ξ, називається достовірною.

Якщо подія напевно не може відбутися при реалізації комплексу умов Ξ, то вона називається неможливою.

Подія A, яка при реалізації комплексу умов Ξ може відбутися, а може й не відбутися, називається випадковою.

Із цих визначень ясно, що, говорячи про достовірність, неможливість або випадковість якої-небудь події, завжди будемо мати на увазі її достовірність, неможливість або випадковість відносно деякого певного комплексу умов.

Науковий інтерес має не просте констатування випадковості події, а кількісна оцінка ймовірності. Ця оцінка виражається твердженням виду:

(2). Імовірність того, що при здійсненні комплексу умов Ξ відбувається подія A, дорівнює p.

Закономірності виду (2) називаються ймовірнісними або стохастичними.

Більшість визначень ймовірності може бути поділена на чотири групи:

1. Означення математичної ймовірності як кількісної міри "ступеня переконаності" дослідника. Це так зване суб'єктивне поняття ймовірності.

2. Означення, що зводять поняття ймовірності до поняття "рівноможливості" (так зване класичне означення ймовірності).

3. Означення, для яких початковим є поняття "частоти" появи події у великій кількості випробовувань (статистичне (частотне) означення ймовірності).

4. Аксіоматичне означення ймовірності.

Ідея про те, що ймовірність випадкової події A за відомих умов допускає кількісну оцінку за допомогою деякого числа, знайшла систематичний розвиток уперше в XVII ст. у працях П. Ферма (1601–1665), Б. Паскаля (1623–1662), Х. Гюйгенса (1629–1695) і особливо Я. Бернуллі (1654–1705). Своїми дослідженнями вони поклали початок теорії ймовірностей в сучасному розумінні слова. Із того часу теорія ймовірностей як математична дисципліна безперервно прогресує, збагачуючись новими важливими результатами.

Рекомендована література

1. Гнєденко Б. В. Курс теорії ймовірностей. – К.: ВПЦ "Київський університет 2008.

2. Гихман И. И., Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика. К., 1988.

3. Карташов М.В. Імовірність, процеси, статистика. – К.: ВПЦ "Київський університет 2008.

4. Скороход А.В. Лекції з теорії випадкових процесів.– К.: Либідь, 1990.

5. Голомозий В.В., Карташов М. В., Ральченко К. В. "Збірник задач з теорії ймовірностей". ВПЦ 'Київський університет', - 2016

1 Випадкові події

1.1 Стохастичний експеримент, випадкові події та операції над ними

Теорія ймовірностей, як будь-який інший розділ математики, побудована аксіоматично, тобто має певні основні вихідні поняття, які не визначаються в рамках теорії, та аксіоми, що пов'язують ці поняття. Основними поняттями теорії ймовірностей є поняття стохастичного експерименту, елементарної події та простору елементарних подій.

Стохастичний експеримент – це експеримент чи спостереження, результат якого не можна передбачити однозначно (наприклад, підкидання монети чи грального кубика). Певний фіксований результат стохастичного експерименту, який не можна виразити через сукупність інших результатів, називається елементарною подією.

Зокрема, різні елементарні події не можуть відбутися одночасно. Множина всіх елементарних подій називається простором елементарних подій і позначається символом грецького алфавіту Ω.

Отже, в результаті проведення стохастичного експерименту завжди відбувається одна і тільки одна елементарна подія яка позначається через ω з простору елементарних подій Ω.

1.1.1 Приклади стохастичних експериментів та просторів елементарних подій

Приклади стохастичних експериментів та просторів елементарних подій

Приклад 1.1. Підкидання однієї монети. Простір елементарних подій має вигляд $\Omega = \{A, P\}$, оскільки результат підкидання, що відрізняється від аверса A, або реверса P, вважається неможливим.

Приклад 1.2. Підкидання двох монет. Оскільки фіксація результату підкидання двох монет потребує задання впорядкованої пари з двох елементів - результатів підкидання першої та другої монети, то простір елементарних подій містить пари можливих варіантів підкидань однієї монети:

$$\Omega = \{AA, AP, PA, PP\}.$$

Приклад 1.3. Підкидання грального кубика. Оскільки гральний кубик має шість граней, а падіння на ребро вважається неможливим,

то
$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

У наведених прикладах простір елементарних подій містить скінченну кількість елементарних подій. Наведемо приклади стохастичних експериментів у яких просторів елементарних подій нескінченний.

Приклад 1.4. Проведення експериментів (підкидання монети, наприклад) до першого успіху. Можна припустити, що успіх обов'язково має відбутися за якоїсь скінченної кількості підкидань, а тому
$$\Omega = \{\text{У}, \text{НУ}, \text{ННУ}, \ldots\}.$$

Приклад 1.5. Вибір випадкової точки на відрізку $[a, b]$. У цьому випадку $\Omega = \{x \in \mathbb{R}\colon a \leq x \leq b\}$, оскільки кожна така точка однозначно задається своєю координатою з відрізку $[a, b]$.

Приклад 1.6. Задача про зустріч. Дві особи A і B домовилися зустрітися в певному місці в інтервалі часу $[0, T]$. Якщо час приходу першої особи протягом указаної години позначити через x, а час приходу другої особи позначити через y, то матимемо
$$\Omega = \{(x, y) \in \mathbb{R}\colon 0 \leq x \leq T, 0 \leq y \leq T\}.$$

1.1.2 Випадкові події –підмножини простору елементарних подій

Зі стохастичним експериментом можна пов'язати певне висловлювання про його результат. Оскільки для кожної елементарної події можна встановити, справедливе дане висловлювання чи ні, то довільному висловлюванню відповідає певна множина елементарних подій, а саме: така підмножина простору елементарних подій, для елементів якої справджується дане висловлювання. Випадковими подіями називаються підмножини простору елементарних подій, що є відповідними прообразами висловлювань про результат стохастичного експерименту.

1.1.3 Приклади випадкових подій

Приклад 1.7. Підкидання двох монет. Позначимо через B випадкову подію {при підкиданні двох монет першим випаде A}, а через через C позначимо випадкову подію {принаймі один раз випаде

$A\}$. Простір всіх елементарних подій $\Omega = \{AA, AP, PA, PP\}$. Події B відповідає підмножина $B = \{AA, AP\}$ простору Ω. Події C відповідає підмножина $C = \{AA, AP, PA\}$ простору Ω.

Приклад 1.8. Підкидання грального кубика. Позначимо через B випадкову подію {при підкиданні грального кубика випаде парна кількість очок}, а через через C позначимо випадкову подію {при підкиданні грального кубика випаде непарна кількість очок}. Простір всіх елементарних подій $\Omega = \{1, 2, 3, 4, 5, 6\}$. Події B відповідає підмножина $B = \{2, 4, 6\}$ простору Ω. Події C відповідає підмножина $C = \{1, 3, 5\}$ простору Ω.

Приклад 1.9. Проведення експериментів до першого успіху. Позначимо через A випадкову подію {успіх настане непізніше четвертого експерименту}, Простір всіх елементарних подій $\Omega = \{\text{У}, \text{НУ}, \text{ННУ}, \ldots\}$. Події A відповідає підмножина $A = \{\text{У}, \text{НУ}, \text{ННУ}, \text{НННУ}\}$ простору Ω.

Приклад 1.10. Вибір випадкової точки на відрізку $[a, b]$. У цьому випадку $\Omega = \{x \in \mathbb{R} \colon a \leq x \leq b\}$. Клас усіх випадкових подій має містити всі інтервали $[x, y] \subset [a, b]$ та їх перетворення.

1.1.4 Властивості класу випадкових подій

Клас усіх випадкових подій у даному стохастичному експерименті будемо позначати через $\mathfrak{F}$.

Оскільки клас висловлювань стосовно результатів експерименту замкнутий відносно операцій заперечення, кон'юнкції, диз'юнкції та інших логічних операцій, то природно описати відповідні властивості класу $\mathfrak{F}$ всіх випадкових подій – підмножин Ω.

1. Множина $\Omega \subset \Omega$ ("щось та відбудеться") є випадковою подією: $\Omega = \{\omega \colon \omega \in \Omega\} \in \mathfrak{F}$, яка називається універсальною подією. Оскільки інші результати експерименту крім тих, які описані точками $\{\omega \colon \omega \in \Omega\}$, відбутись не можуть, то множину Ω називають ще достовірною подією.

2. Множина $\varnothing$, що не містить жодної елементарної події ("нічого не відбудеться"), є випадковою подією: $\varnothing = \{\omega \colon \omega \notin \Omega\} \in \mathfrak{F}$, і називається неможливою подією.

3. Належність елементарної події ω випадковій події A, позначення $\omega \in A$, відображається висловлюваннями: ω сприяє A, або ж подія A відбувається при даній елементарній події ω.

4. Справедливість включення $A \subset B$ визначає, що подія A спричиняє подію B (або ж міститься в ній). Альтернативна інтерпретація: якщо відбувається A, то відбувається B. Очевидно що для будь-якого A маємо $A \subset \Omega$. За означенням приймають, що $\varnothing \subset A$.

5. Для події A її доповнення (або заперечення)

$$\overline{A} = \{\omega \colon \omega \notin A\} \in \mathfrak{F},$$

є випадковою подією, яка полягає в тому, що не відбудеться A. Справджуються такі співвідношення

$$A \cup \overline{A} = \Omega, \quad A \cap \overline{A} = \varnothing.$$

6. Для двох подій A та B їх об'єднання

$$A \cup B = \{\omega \colon \omega \in A \text{ або } \omega \in B\} \in \mathfrak{F}$$

є випадковою подією, яка полягає в тому, що відбудеться або A, або B, або обидві події разом. Справджуються співвідношення

$$A \cup \Omega = \Omega, \quad A \cup \varnothing = A.$$

Подія $A \cup B$ грає роль точної верхньої грані подій A та B. Це потрібно розуміти так: по перше

$$A \subset A \cup B, \quad B \subset A \cup B,$$

а якщо подія C така, що $A \subset C$ та $B \subset C$, то $A \cup B \subset C$.

7. Для двох випадкових подій A та B їх переріз (або ж перетин) – множина

$$A \cap B = \{\omega \colon \omega \in A \ \text{ та } \ \omega \in B\} \in \mathfrak{F}$$

є випадковою подією, яка полягає в тому, що події A та B відбудуться одночасно.

Відмітимо співвідношення

$$A \cap \Omega = A, \quad A \cap \varnothing = \varnothing,$$

$$A \cap B \subset A, \quad A \cap B \subset B,$$

а якщо подія C така, що $C \subset A \cap B$, то $C \subset A$, а також $C \subset B$.

8. Якщо $A \cap B = \varnothing$, то події A та B називаються несумісними (або ж такими, що не перетинаються).

9. Для двох подій A та B їх різниця

$$A \setminus B = \{\omega \colon \omega \in A \text{ та } \omega \notin B\} \in \mathfrak{F}$$

є випадковою подією, яка полягає в тому, що відбудеться A і одночасно не відбудеться B.

Відмітимо співвідношення

$$\overline{A} = \Omega \setminus A, \quad A \setminus B = A \cap \overline{B}.$$

10. Для двох подій A та B їх симетрична різниця $A \triangle B$ є випадковою подією, яка полягає в тому, що з подій A та B відбудеться точно одна подія, тобто

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

11. Події послідовності $\{A_n, n \geq 1\}$ називаються попарно несумісними, якщо $A_i \cap A_j = \varnothing$ для всіх $i \neq j$.

12. Для послідовності випадкових подій $\{A_n, n \geq 1\}$ їх зліченне об'єднання

$$\cup_{n \geq 1} A_n = \{\omega \colon \exists n \geq 1 \colon \omega \in A_n\} \in \mathfrak{F},$$

є випадковою подією, яка полягає в тому, що відбудеться хоча б одна подія із послідовності $\{A_n, n \geq 1\}$.

13. Для послідовності випадкових подій $\{A_n, n \geq 1\}$ їх злічений переріз

$$\cap_{n \geq 1} A_n = \{\omega \colon \omega \in A_n, \forall n \geq 1\} \in \mathfrak{F}$$

є випадковою по дією, яка полягає в тому, що відбудуться всі події із даної послідовності $\{A_n, n \geq 1\}$.

14. Нехай $\{A_n, n \geq 1\}$ – монотонно неспадна послідовність подій, тобто $A_n \subset A_{n+1}, \forall n \geq 1$. Їх монотонною границею A (що позначається як $A_n \uparrow A = \lim A_n$ є випадкова подія, яка полягає в тому, що відбудеться принаймні одна подія даної послідовності (а отже, і кожна, починаючи з деякого номера), тобто $A = \cup A_n$. У такому випадку кажуть, що події A_n монотонно збігаються до A.

15. Нехай $\{A_n, n \geq 1\}$ – монотонно незростаюча послідовність подій, тобто $A_{n+1} \subset A_n, \forall n \geq 1$. Їх монотонною границею А є випадкова подія, яка полягає в тому, що відбудуться одночасно всі події послідовності, тобто $A = \cap A_n$ (це позначається як $A_n \downarrow A = \lim A_n$). У такому випадку кажуть, що події A_n монотонно збігаються до A.

16. Для послідовності подій $\{A_n, n \geq 1\}$ їх верхня границя $\overline{\lim} A_n$ є випадковою подією, яка полягає в тому, що події даної послідовності відбудуться нескінченно часто, тобто відбудуться всі події з деякої нескінченної підпослідовності A_{n_k}. Цю подію можна зобразити у вигляді $\overline{\lim} A_n = \cap_{m \geq 1} \cup_{n \geq m} A_n$. Дійсно, елементарна подія ω належить правій частині рівності тоді й тільки тоді, коли для довільного номера $m \geq 1$ знайдеться номер $n \geq m$ такий, що $\omega \in A_n$, тобто коли відбудеться нескінченна кількість серед подій A_n.

17. Для послідовності подій $\{A_n, n \geq 1\}$ їх нижня границя $\underline{\lim} A_n$ є випадковою подією, яка полягає в тому, що відбудуться всі події даної послідовності починаючи з деякого номера. Ця подія позначається через $\overline{\lim} A_n = \cup_{m \geq 1} \cap_{n \geq m} A_n$. Дійсно, елементарна подія ω належить правій частині рівності тоді й тільки тоді, коли знайдеться номер $m \geq 1$ такий, що $\omega \in A_n$ при всіх $n \geq m$, тобто коли відбудуться всі події A_n, починаючи з деякого номера.

Між теоретико-ймовірнісними поняттями випадкових подій і операцій над ними, та теоретико-множинними поняттями множин і їх перетворень, є пряма відповідність, що ілюструється у наступній таблиці.

Вправа 1.1. Довести, що: (a) $A \setminus B = A \setminus (A \cap B)$, (b) $A \triangle B = (A \cup B) \setminus (A \cap B)$, (c) $(A \triangle B) \triangle C = A \triangle (B \triangle C)$, (d) $\overline{A \cup B} = \overline{A} \cap \overline{B}$, (e) $\overline{A \triangle B} = \overline{A} \triangle \overline{B}$, (f) $(A \cap B) \cap C = A \cap (B \cap C)$, (g) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

Познач.	У теорії ймовірностей	У теорії множин
Ω	Простір елементарних подій	Універсальна множина
ω	Елементарна подія	Елемент множини
$\varnothing$	Неможлива подія	Порожня множина
A	Випадкова подія	Множина
$\omega \in A$	ω сприяє події A	ω належить A
$\overline{A}$	Заперечення події A	Доповнення множини A
$A \subset B$	Подія A спричиняє B	A міститься в B
$A \cap B$	Відбудуться події A та B	Переріз множин A і B
$\cap_{n\geq 1} A_n$	Відбудуться всі події A_n	Переріз множин A_n
$A \cap B = \varnothing$	Події A і B несумісні	A і B не перетинаються
$A \cup B$	Відб. подія A або B	Об'єднання A і B
$\cup_{n\geq 1} A_n$	Відб. хоча б одна з подій A_n	Об'єднання множин A_n
$A \setminus B$	Відб. подія A, та не B	Різниця множин A та B
$A \triangle B$	Відб. точно одна з A, B	Симетр. різниця A, B
$A_n \uparrow A$	A_n не спадають та $\exists \lim A_n$	$A_n \subset A_{n+1}$, $A = \cup A_n$
$A_n \downarrow A$	A_n не зростають та $\exists \lim A_n$	$A_{n+1} \subset A_n$, $A = \cap A_n$
$\overline{\lim} A_n$	A_n відбув. нескінченно часто	$\cap_{m\geq 1} \cup_{n\geq m} A_n$
$\underline{\lim} A_n$	A_n відбув. всі, з деякого m	$\cup_{m\geq 1} \cap_{n\geq m} A_n$

Таблиця 1.1

1.2 Основи комбінаторики

1.2.1 Основний принцип комбінаторики. Правило множення

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**.

Щоб сформулювати це правило, розпочнемо з такої задачі, яка яскраво ілюструє ситуацію.

Приклад 1.11. З міста A до міста B веде n *різних* доріг, з міста B до міста C веде m різних шляхів. Скількома способами можна здійснити подорож з міста A в місто C?

Відповідь очевидна: $n \cdot m$, бо обравши один з n можливих шляхів з міста A в місто B, ми матимемо m різних продовжень подорожі з B в C.

Міркування, які були проведені при розв'язанні, задачі доводять справедливість такого простого правила, яке називають **правилом множення** або **основним принципом комбінаторики**.

Якщо деякий вибір A можна здійснити n способами, а для кожного з цих способів другий вибір можна здійснити m способами, то вибір A і B (у вказаному порядку) можна здійснити $n \cdot m$ способами.

Приклад 1.12. На вершину гори веде 8 доріг.

а) Скількома способами турист може піднятися на гору і спуститися вниз?

б) Скількома способами турист може піднятися на гору і спуститися вниз, якщо підйом і спуск здійснюються різними шляхами?

а) Відповідь: $8 \times 8 = 64$. б)Відповідь: $8 \times 7 = 56$.

Сформулюємо тепер **правило множення в загальному вигляді**.

Нехай треба виконати одна за одною k дій. Якщо першу дію можна здійснити n_1 способами, другу дію (після того, як здійснена перша) — n_2 способами, третю дію (після того, як здійснені попередні дві) — n_3 способами, і так до k-тої дії, яку можна здійснити n_k способами, то загальне число способів здійснення дій дорівнює

$$n_1 \times n_2 \times \cdots \times n_k.$$

Користуючись правилом множення неважко переконатися в справедливості такої теореми, яку часто використовують в теорії чисел.

Теорема 1.1. *Нехай розклад натурального числа n на прості множники має вигляд*

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \dots p_k^{\alpha_k}.$$

Тоді число усіх дільників числа n дорівнює

$$(\alpha_1 + 1)(\alpha_2 + 1)\dots(\alpha_k + 1).$$

Доведення. Кожен дільник числа n має вигляд

$$p_1^{i_1} \cdot p_2^{i_2} \dots p_k^{i_k},$$

де $0 \le i_1 \le \alpha_1$, $0 \le i_2 \le \alpha_2, \dots, 0 \le i_k \le \alpha_k$.

Обрати дільник — це обрати i_1 ($(\alpha_1 + 1)$ можливих значень), i_2 ($(\alpha_2 + 1)$ можливих значень),..., обрати i_k ($(\alpha_k + 1)$ можливих значень). Для завершення доведення теореми досить застосувати правило множення. $\square$

Користуючись правилом множення можна підрахувати число відображень множини X в множину Y.

Означення 1.1. Будемо говорити, що задано відображення $\varphi(x) : X \to Y$ множини X в множину Y, якщо кожному x з множини X поставлено у відповідність y з множини Y.

Теорема 1.2. *Нехай X та Y скінченні множини, $N(X) = n$, $N(Y) = m$. Тоді число усіх відображень X в Y дорівнює m^n.*

1.2.2 Число підмножин $n-$елементної множини

Нагадаємо, якщо *кожен* елемент x, який належить множині A, належить також і множині B, то множина A називається *підмножиною* множини B; позначається це так: $A \subset B$.

Поставимо таке питання. Нехай множина X містить n елементів, $N(X) = n$. Скільки *різних* підмножин має множина X?

Домовимося вважати, що порожня множина $\emptyset$ є підмножиною будь-якої множини.

Теорема 1.3. *Число усіх підмножин множини X, яка містить n елементів, дорівнює 2^n.*

Доведення. Нехай $A \subset X$. Поставимо у відповідність A набір з нулів та одиниць за таким правилом: на k-тому місті цього набору пишемо 1, якщо елемент x_k належить множині A, і нуль, якщо елемент x_k не належить множині A. Отже, кожній підмножині відповідає свій набір нулів та одиниць. Наприклад, порожній підмножині відповідатиме набір складений з одних нулів; підмножині $\{x_1, x_n\}$, що містить два елементи з X, відповідатиме набір $(1, 0, \dots, 1)$. Відповідність між підмножинами та наборами з нулів та одиниць є взаємно однозначною. Число всіх можливих наборів, які можна скласти з нулів та одиниць, дорівнює згідно з правилом множення

$$\underbrace{2 \times 2 \times \dots \times 2}_{n-\text{раз}} = 2^n.$$

Тому й число усіх підмножин дорівнює 2^n. $\square$

Множину усіх підмножин множини X називають булеаном множини на честь англійського математика Джорджа Буля (1815-1864). Позначають булеан так: 2^X. Ми встановили таким чином, що $N(2^X) = 2^n$. Іншими словами $N(2^X) = 2^{N(X)}$.

1.2.3 Число k-елементних підмножин множини X

Наступне питання. Скільки серед всіх підмножин множини X є k-елементних підмножин (підмножин, які містять k елементів)?

Позначимо через C_n^k число k-елементних підмножин множини, яка містить n елементів. Без обчислень можна зазначити, що $C_n^1 = n$, $C_n^k = C_n^{n-k}$ (кожній k-елементній підмножині ми можемо поставити у відповідність її доповнення, і ця відповідність буде взаємно однозначною відповідністю).

Теорема 1.4. *Число всіх k-елементних підмножин множини, що складається з n елементів, дорівнює*

$$C_n^k = \frac{n(n-1)\dots(n-k+1)}{1 \cdot 2 \cdot 3 \dots k} = \frac{n!}{k!(n-k)!}\ .$$

Зазначимо, що символ $n!$ (читається n-факторіал) позначає добуток $1 \cdot 2 \cdot 3 \dots n$ всіх натуральних чисел від 1 до n включно. Зручно вважати, що $0! = 1$.

Доведення. Щоб побудувати k-елементну підмножину множини X треба до $k-1$-елементної підмножини додати один з $n-(k-1)$ елементів, які не входять в цю підмножину. Оскільки число $(k-1)$-елементних підмножин дорівнює C_n^{k-1}, і кожну з них можна перетворити в k-елементну $n-k+1$ способом, то таким чином, ми дістанемо $(n-k+1)C_n^{k-1}$ підмножин. Але не всі вони будуть різними, бо кожну k-елементну підмножину можна побудувати k способами:

додаванням кожного з її елементів. Тому обчислене нами число в k раз більше, ніж число k-елементних підмножин. Отже,

$$kC_n^k = (n-k+1)C_n^{k-1},$$

звідки отримуємо рекурентне співвідношення

$$C_n^k = \frac{(n-k+1)}{k} C_n^{k-1}.$$

Повторимо (проітеруємо) це співвідношення

$$C_n^k = \frac{(n-k+1)}{k} C_n^{k-1} = \frac{(n-k+1)}{k} \frac{(n-k+2)}{k-1} C_n^{k-2} = \cdots =$$

$$= \frac{(n-k+1)}{k} \frac{(n-k+2)}{k-1} \cdots \frac{(n-k+(k-1))}{k-(k-2)} C_n^1 =$$

$$= \frac{(n-k+1)(n-k+2)\ldots(n-1)n}{k(k-1)\ldots 2\cdot 1}\ .$$

Домноживши чисельника і знаменника на $(n-k)!$, отримаємо

$$C_n^k = \frac{n!}{k!(n-k)!}. \tag{1.1}$$

□

Числа C_n^k називаються *біноміальними коефіцієнтами*. Пояснення цієї назви буде дано в наступному параграфі. Довільна k-елементна підмножина n-елементної множини називається комбінацією елементів з n по k.

В підручниках з комбінаторики, які не вживають теоретико-множинної термінології, дають таке означення.

Комбінація з n по k — група k предметів з даних n предметів, причому порядок розміщення предметів в групі не є істотним.

Приклад 1.13. Довести, використовуючи лише комбінаторні міркування (тобто не використовуючи (1.1)), що

$$C_n^k = C_{n-1}^{k-1} + C_{n-1}^k. \tag{1.2}$$

Розв'язання. Нехай $X = \{x_1, x_2, \ldots, x_n\}$. Зафіксуємо який-небудь елемент, наприклад x_1. Розіб'ємо всі k-елементні підмножини на два класи $\mathcal{K}_1$ і $\mathcal{K}_2$. До $\mathcal{K}_1$ віднесемо ті k-елементні підмножини, які містять в собі елемент x_1; до $\mathcal{K}_2$ віднесемо ті k-елементні підмножини, які не містять x_1. Число підмножин, які входять до $\mathcal{K}_2$, дорівнює C_{n-1}^k, бо кожна така множина є k-елементною підмножиною $(n-1)$-елементної множини $\{x_2, x_3, \ldots, x_n\}$. Число підмножин, які входять до $\mathcal{K}_1$, дорівнює C_{n-1}^{k-1}, бо, викинувши з кожної такої множини x_1 одержимо $(k-1)$-елементну підмножину множини $\{x_2, \ldots, x_n\}$. Отже, $C_n^k = C_{n-1}^{k-1} + C_{n-1}^k$.

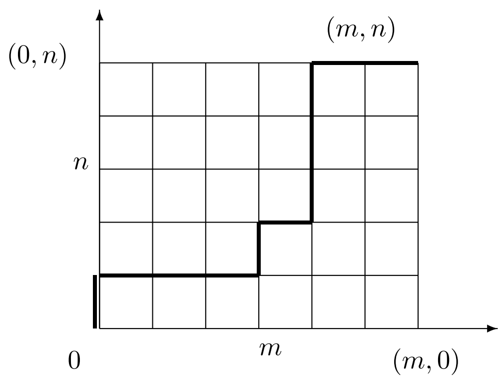
Приклад 1.14. Довести рівність

$$\sum_{k=0}^{n} C_n^k = 2^n. \tag{1.3}$$

Розв'язання. Оскільки C_n^k — число k-елементних підмножин множини, яка містить n елементів, то сума, яка стоїть в лівій частині рівності (1.3), являє собою число усіх підмножин n-елементної множини ($C_n^0 = 1$ відповідає порожній множині).

Геометрична інтерпретація C_n^k.

Розглянемо прямокутну сітку квадратів розмірами $m \times n$ ("шахове місто яке складається з $m \times n$ прямокутних кварталів, поділених $n-1$ "горизонтальними" і $m-1$ "вертикальними" вулицями (рис. 7)). Яке число різних найкоротших доріг на цій сітці, які ведуть з лівого нижнього кута (точки $(0,0)$) в правий верхній кут (точку (m,n))?



(Рис. 7)

Відзначимо, що кожен найкоротший шлях з точки $(0,0)$ в точку (m,n) складається з $m+n$ відрізків, причому серед них є m горизонтальних і n вертикальних відрізків. Обравши вертикальні відрізки, ми повністю визначаємо шлях. Тому загальне число шляхів дорівнює числу способів, якими з $m+n$ відрізків можна вибрати n вертикальних відрізків, тобто C_{m+n}^n.

Можна було розглядати вибір не n вертикальних, а m горизонтальних відрізків, і ми отримали б тоді відповідь C_{m+n}^m. Таким чином, ми встановили рівність $C_{m+n}^n = C_{m+n}^m$, в справедливості якої неважко переконатися, використовуючи рівність (1.1), і довели таке твердження.

Теорема 1.5. *Число найкоротших шляхів з точки* $(0,0)$ *в точку* (m,n) *дорівнює*

$$C_{m+n}^m = C_{m+n}^n.$$

1.2.4 Впорядковані множини. Перестановки та розміщення

Означення 1.2. 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.

Кожну скінченну множину можна перетворити у впорядковану, якщо, наприклад, переписати всі елементи множини в деякий список, а потім поставити у відповідність кожному елементу номер місця, на якому він стоїть в списку.

Теорема 1.6. *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. За правилом множення всі n місць можна заповнити способами $n(n-1)\cdots 2\cdot 1$. Отже, множину A, яка містить n елементів можна впорядкувати $n!$ способами. $\square$

Різні впорядкування певної множини називають *перестановками* цієї множини.

Приклад 1.15. Скількома способами можна впорядкувати множину $\{1, 2, 3, \ldots, 2n\}$ так, щоб кожне парне число мало парний номер?

Розв'язання. Парні числа можна розставити на місцях з парними номерами (таких місць n) $n!$ способами; кожному способу розташування парних чисел на місцях з парними номерами відповідає $n!$ способів розташування непарних чисел на місцях з непарними номерами. Тому загальне число перестановок вказаного типу згідно з правилом множення дорівнює $(n!) \times (n!) = (n!)^2$.

1.2.5 Впорядковані підмножини даної множини (розміщення)

Нехай множина A містить n елементів. Скільки різних впорядкованих k-елементних підмножин можна утворити з її елементів?

Приклад 1.16. Перерахуємо усі двохелементні впорядковані підмножини трьохелементної множини $\{a, b, c\}$: $\{a, b\}$, $\{b, a\}$, $\{a, c\}$, $\{c, a\}$, $\{b, c\}$, $\{c, b\}$.

Означення 1.3. Впорядковані k-елементні підмножини називають розміщеннями з n по k.

В підручниках з комбінаторики, які не використовують термінологію з теорії множин, дають таке означення розміщень.

Означення 1.4. Розміщення з n по k — це групи по k предметів із даних n предметів, причому порядок предметів в групі є істотним.

В прикладі ми перерахували всі розміщення з 3 по 2.

Теорема 1.7. *Число розміщень з n по k дорівнює*

$$A_n^k = n(n-1)\ldots(n-k+1). \tag{1.4}$$

Доведення. Будемо конструювати k-елементну підмножину шляхом послідовного вибору елементів: перший елемент може бути обраний n способами, другий — $(n-1)$ способами, після того як уже обрано $(k-1)$ елементів, наступний k-ий елемент може бути обраний $n-(k-1)=n-k+1$ способами. Тому згідно з правилом множення $A_n^k = n(n-1)\dots(n-k+1)$. $\square$

Зазначимо, що в англомовній літературі для числа розміщень з n по k вживається таке позначення

$$(n)_k = n(n-1)\cdots(n-k+1).$$

Теорема 1.8. *Має місце співвідношення між числом розміщень A_n^k із n по k і числом комбінацій C_n^k з n по k*

$$A_n^k = C_n^k \cdot k!. \tag{1.5}$$

Доведення. Розглянемо всі k-елементні підмножини множини A (їх число позначаємо C_n^k). Кожну з них упорядкуємо (це можна зробити $k!$ способами). Тоді ми одержимо всі k-елементні впорядковані підмножини. Отже, $A_n^k = C_n^k \cdot k!$. $\square$

Таким чином, ми отримали ще одне зовсім коротке доведення формули для C_n^k:

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

Приклад 1.17. Студенту треба скласти 4 екзамени впродовж 8 днів. Скількома способами це можна зробити?

Розв'язання. Шукане число способів дорівнює числу 4-елементних впорядкованих підмножин (дні складання екзаменів) множини з 8 елементів, тобто $A_8^4 = 8\cdot 7\cdot 6\cdot 5 = 1680$ способів. Якщо студент хоче складати останній екзамен на восьмий день, то число способів дорівнює $4\cdot A_7^3 = 4\cdot 7\cdot 6\cdot 5 = 840$.

1.2.6 Перестановки з повтореннями

Розглянемо множину A, яка містить n елементів. Поставимо таке питання: Скількома способами можна представити множину A у вигляді об'єднання множин B_1 і B_2 таких, що $N(B_1)=k_1$, $N(B_2)=k_2$, $(B_1\cap B_2=\emptyset;\ k_1+k_2=n)$?

Множину B_1 ми можемо обрати $C_n^{k_1}$ способами; до множини B_2 зараховуємо $n - k_1 = k_2$ елементів, які залишилися після вибору B_2. Таким чином, шукане число способів дорівнює

$$C_n^{k_1} = \frac{n!}{k_1!(n-k_1)!} = \frac{n!}{k_1!k_2!}.$$

Наступне питання. Скількома способами можна представити множину A у вигляді об'єднання m підмножин

$$A = B_1 \cup B_2 \cup \cdots \cup B_m$$

так, щоб $N(B_1) = k_1$, $N(B_2) = k_2$,...,$N(B_m) = k_m$, де $k_1, k_2, \ldots, k_m$ дані числа $(k_1 \geq 0, k_2 \geq 0, \ldots, k_m \geq 0; k_1 + k_2 + \cdots + k_m = n)$? Множини $B_1, B_2, \ldots, B_m$ не повинні мати спільних елементів. Всі описані вище розбиття множини A на m груп можна побудувати так: візьмемо довільну k_1-елементну підмножину B_1 множини A (це можна зробити $C_n^{k_1}$ способами); з $n - k_1$ елементів, які залишились, утворимо k_2-елементну підмножину B_2 (це можна зробити $C_{n-k_1}^{k_2}$ способами) і т.д. Загальне число способів вибору m множин за правилом множення дорівнює

$$C_n^{k_1} C_{n-k_1}^{k_2} C_{n-k_1-k_2}^{k_3} \ldots C_{n-k_1-k_2-\cdots-k_m}^{k_m} =$$

$$= \frac{n!}{k_1!(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \cdot \frac{(n-k_1-k_2)!}{k_3!(n-k_1-k_2-k_3)!} \times \cdots$$

$$\times \frac{(n-k_1-k_2-\cdots-k_{m-1})!}{k_m!(n-k_1-k_2-\cdots-k_m)!} \cdot = \frac{n!}{k_1!k_2!\ldots k_m!}$$

(нагадаємо, що $= 0! = 1$).

Отже, ми довели таку теорему.

Теорема 1.9. *Нехай $k_1, k_2, \ldots, k_m$ — цілі невід'ємні числа, причому $k_1 + k_2 + \cdots + k_m = n$. Число способів, якими можна представити множину A, яка містить n елементів, у вигляді об'єднання множин, число елементів яких становить відповідно $k_1, \ldots, k_m$, дорівнює*

$$C_n(k_1, \ldots k_m) = \frac{n!}{k_1!k_2!\ldots k_m!}. \quad (1.6)$$

Числа $C_n(k_1, \dots k_m)$ називаються поліноміальними коефіцієнтами. В англомовній літературі вони позначаються так

$$\binom{n}{k_1, k_2, \dots k_m}.$$

Ці числа мають ще одну комбінаторну інтерпретацію.

Нехай маємо k_1 букв a_1, k_2 букв a_2, $\dots, k_m$ букв a_m, $k_1 + k_2 + \dots + k_m = n$. Підрахуємо, скільки різних слів можна утворити з цих букв. Перенумеруємо місця, на яких стоять букви, числами $1, 2, \dots, n$. Кожне слово визначається множинами B_1 (номери місць, де стоїть буква a_1), B_2 (номери місць, де стоїть буква a_2), $\dots$, B_m (номери місць, де стоїть буква a_m) . Тому число різних слів дорівнює числу способів, якими можна представити множину $A = \{1, 2, \dots n\}$ у вигляді об'єднання множин $B_1, \dots B_m$, тобто дорівнює

$$C_n(k_1, \dots, k_m) = \frac{n!}{k_1! \dots k_m!}.$$

Приклад 1.18. Число різних слів, які дістанемо, переставляючи літери в слові **математика**, дорівнює $\frac{10!}{2!3!2!} = 151200$.

Звичайно, можна говорити не про букви, а про предмети різних типів. Твердження, сформульоване вище, можна сформулювати так.

Теорема 1.10. *Число різних перестановок предметів, серед яких є k_1 предметів першого типу, k_2 предметів другого типу, $\dots$, k_m предметів m-того типу, дорівнює*

$$C_n(k_1, \dots k_m) = \frac{n!}{k_1! k_2! \dots k_m!}.$$

Наведемо ще одне доведення цієї теореми.

Доведення. Позначимо через $C_n(k_1, \dots k_m)$ число усіх перестановок описаного в умові теореми типу. Розглянемо одну перестановку і замінимо в ній однакові предмети різними. Тоді число усіх перестановок, які можна утворити з розглянутої нами перестановки, дорівнює $k_1! k_2! \dots k_m!$. Зробивши це для кожної перестановки, дістанемо $n!$ перестановок. Отже,

$$C_n(k_1, \dots k_m) k_1! k_2! \dots k_m! = n!,$$

звідки

$$C_n(k_1, \dots k_m) = \frac{n!}{k_1!k_2!\dots k_m!}.$$

$\square$

1.2.7 Комбінації з повтореннями

Означення 1.5. Комбінаціями з m по n з повтореннями будемо називати групи, що містять n предметів, причому кожен предмет належить до одного з m типів, а порядок розташування предметів в групі не є істотним.

Число усіх можливих комбінацій з m по n з повтореннями позначимо f_m^n. Розглянемо декілька прикладів.

Приклад 1.19. Нехай типи предметів a і b, $n = 2$. Тоді комбінації з повторенням мають вигляд aa, ab, bb. Отже, безпосереднім підрахунком переконуємося, що $f_2^2 = 3$.

Приклад 1.20. Нехай типи предметів a і b, $n = 3$. Комбінації з повторенням мають вигляд aaa, aab, abb, bbb. Таким чином, $f_2^3 = 4$.

Приклад 1.21. Нехай типи предметів a, b, c; $m = 3, n = 2$. Комбінації з трьох по два мають вигляд: aa, ab, ac, bb, bc, cc. Отже, $f_3^2 = 6$.

Приклад 1.22. Комбінації з трьох по чотири ($m = 3$; типи a, b, c; $n = 4$) мають вигляд:

$$\begin{array}{ccccc} aaaa & aaab & aaac & aabb & aabc \\ aacc & abbb & abbc & abcc & accc \\ bbbb & bbbbc & bbcc & bccc & cccc\,. \end{array}$$

Таким чином, $f_3^4 = 15$.

Теорема 1.11. *Число усіх комбінацій з m по n з повтореннями дорівнює*

$$f_m^n = C_{n+m-1}^{m-1} = C_{n+m-1}^n. \tag{1.7}$$

Доведення. Кожна комбінація повністю визначається, якщо вказати, скільки предметів кожного з m типів в неї входить. Поставимо у відповідність кожній комбінації набір з нулів та одиниць, утворений за таким правилом: напишемо підряд стільки одиниць, скільки предметів першого типу входить в комбінацію, далі поставимо нуль і після нього напишемо стільки одиниць, скільки предметів другого типу містить ця комбінація і т.д.

Таким чином, кожній комбінації з m по n відповідає набір з n одиниць та $m-1$ нулів і, навпаки, за кожним таким набором однозначно відновлюється комбінація. Тому число комбінацій з m по n з повтореннями дорівнює числу наборів довжини $n+m-1$, які складаються з n одиниць та $m-1$ нулів, тобто дорівнює

$$C_{n+m-1}^{m-1} = C_{n+m-1}^{n}.$$

□

Проілюструємо алгоритм для комбінацій з прикладів 1.19 і 1.20.
Приклад 1.19. $(n=3, m=2)$

$$\begin{array}{ccc} aaa & \longleftrightarrow & 1110 \\ aab & \longleftrightarrow & 1101 \\ abb & \longleftrightarrow & 1011 \\ bbb & \longleftrightarrow & 0111 \end{array}.$$

Приклад 1.20. $(n=2, m=2)$

$$\begin{array}{ccc} aa & \longleftrightarrow & 1100 \\ ab & \longleftrightarrow & 1010 \\ ac & \longleftrightarrow & 1001 \\ bb & \longleftrightarrow & 0110 \\ bc & \longleftrightarrow & 0101 \\ cc & \longleftrightarrow & 0011 \end{array}.$$

Приклад 1.23. Кості з доміно є комбінаціями з повтореннями з семи цифр 0,1,2,3,4,5,6 по 2. Число усіх таких комбінацій дорівнює

$$f_7^2 = C_{2+7-1}^2 = C_8^6 = C_8^2 = 28.$$

Розглянемо ще декілька застосувань теореми 3.

Теорема 1.12. *Число цілих невід'ємних розв'язків рівняння*

$$x_1 + x_2 + \cdots + x_m = n \tag{1.8}$$

дорівнює $f_m^n = C_{n+m-1}^{m-1} = C_{n+m-1}^{n}$.

Доведення. Існує взаємно однозначна відповідність між розв'язками рівняння (1.8) та комбінаціями з m предметів по n. Якщо маємо невід'ємні цілі числа $x_1, x_2, \ldots, x_n$, які задовольняють рівнянню (1.8), то можемо утворити комбінацію з m по n, взявши x_1 предметів першого типу, x_2 предметів другого типу,. . . , x_m m- того типу. Навпаки, маючи комбінацію з m предметів по n з повтореннями, дістанемо розв'язок рівняння (1.8) (x_1 — число предметів першого типу, x_2 — число предметів другого типу,..., x_m — число предметів m-того типу). Тому число цілих невід'ємних розв'язків рівняння (3) дорівнює

$$f_m^n = C_{n+m-1}^{m-1} = C_{n+m-1}^n.$$

□

Теорема 1.13. *Число цілих додатних розв'язків рівняння*

$$x_1 + x_2 + \cdots + x_m = n \tag{1.9}$$

$$(x_1 > 0, x_2 > 0, \ldots, x_m > 0)$$

дорівнює C_{n-1}^{m-1}.

Доведення. Зазначимо, що $m \leq n$ є необхідною умовою існування розв'язку рівняння (1.9).

Віднімемо від обох частин рівняння (1.9) m. Отримаємо

$$(x_1 - 1) + (x_2 - 1) + \cdots + (x_m - 1) = n - m.$$

Покладемо $x_i - 1 = y_i$. Тоді $y_i \geq 0$ і $y_1, \ldots, y_m$ є невід'ємними розв'язками рівняння

$$y_1 + y_2 + \cdots + y_m = n - m.$$

Згідно з попередньою теоремою число невід'ємних розв'язків цього рівняння дорівнює

$$C_{(n-m)+m-1}^{m-1} = C_{n-1}^{m-1}.$$

□

Приклад 1.24. Скількома способами можна вибрати 3 букви з 12 букв А, А, А, Т, Т, Т, Г, Г, Г, Ц, Ц, Ц?

Розв'язання. Маємо справу з комбінаціями з повтореннями: $m = 4$ (чотири типи предметів А, Т, Г, Ц), а $n = 3$. Тому шукане число $f_4^3 = C_{3+4-1}^3 = C_6^3 = 20$.

Задача має відношення до теорії білкового коду, запропонованої американським фізиком російського походження Г. Гамовим. Букви А, Т, Г, Ц позначають нуклеотиди: аденін, тимін, гуанін, цитозин. Число трійок нуклеотидів дорівнює 20 — числу стандартних амінокислот, на які розкладається молекула білка.

Приклад 1.25. Скільки різних часткових похідних порядка n має нескінченно диференційована функція m змінних?

Розв'язання. Часткові похідні порядка n нескінченно диференційованої функції не залежать від порядку диференціювання, а залежать лише від того, скільки раз ми диференціюємо по кожній змінній. Якщо за першою змінною диференціювати k_1 раз, за другою — k_2 раз, ..., за m-ою k_m раз ($k_1+k_2+\cdots+k_m = n$), то отримаємо часткову похідну

$$\frac{\partial^n f(x_1, x_2, \ldots, x_m)}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_m^{k_m}}.$$

Тому число різних часткових похідних дорівнює числу невід'ємних розв'язків рівняння (3), тобто дорівнює

$$f_m^n = C_{n+m-1}^{m-1}.$$

Наприклад, функція трьох змінних має 15 різних похідних четвертого порядку.

1.2.8 Розміщення частинок в комірках

Розміщення частинок в комірках (модель Максвела-Больцмана, модель Бозе-Ейнштейна, модель Фермі-Дірака). Розглянемо застосування отриманих вище результатів до питань розміщення n частинок в m комірках. Ці питання грають важливу роль в статистичній фізиці.

Теорема 1.14. *Нехай n частинок, які можна розрізнити, розподілені в m комірках (областях простору). Тоді число всіх можливих розміщень частинок дорівнює m^n. Число таких розміщень, при яких в першій комірці знаходиться k_1 частинок, в другій —*

k_2 частинок,..., в m-ій — k_m частинок дорівнює

$$C_n(k_1, k_2, \dots k_m) = \frac{n!}{k_1! \dots k_m!}.$$

Доведення. Перенумеруємо комірки номерами $1, 2, \dots, m$ і частинки номерами $1, 2, \dots, n$. Розташування частинок визначене, якщо кожній частинці з номером k вказано номер i_k комірки, де вона знаходиться. Таким чином, розташування частинок описується набором $(i_1, i_2, \dots i_n)$, де кожне i_k може приймати одне з m значень $1, 2, \dots, m$. Число таких наборів згідно з правилом множення дорівнює

$$\underbrace{m \cdot m \dots m}_{n-\text{раз}} = m^n.$$

Отже, число усіх можливих розміщень частинок дорівнює m^n.

Нехай B_1 — множина номерів частинок, які потрапили в першу комірку, B_2 — множина номерів частинок з другої комірки,..., B_m — множина номерів частинок з m-ої комірки. Тоді

$$\{1, 2, \dots, n\} = \cup_{i=1}^m B_i.$$

Згідно з Теоремою 1.9 число тих розміщень частинок, при яких в першій комірці k_1 частинок, в другій k_2 частинок,..., в m-ій комірці k_m частинок дорівнює $C_n(k_1, \dots, k_m)$. □

Розглянута нами модель розміщення частинок в статистичній фізиці називається **моделлю Максвела-Больцмана**.

Наслідок 1.1. *Як наслідок маємо таке співвідношення*

$$\sum_{\substack{k_1+\dots+k_m=n \\ k_1\geq 0,\dots,k_m\geq 0}} C_n(k_1, \dots, k_m) = m^n. \tag{1.10}$$

Аналізуємо тепер розміщення однакових частинок.

Теорема 1.15. *Нехай n однакових частинок розміщені в m комірках. Тоді число всіх можливих розміщень частинок дорівнює*

$$f_m^n = C_{n+m-1}^{m-1} = C_{n+m-1}^n.$$

Число тих розміщень, при яких в кожній комірці є принаймні одна частинка, дорівнює

$$C_{n-1}^{m-1}.$$

Доведення. Оскільки частинки не розрізняються, число розміщень частинок визначається набором $(x_1, x_2, \dots x_m)$, де x_i — число частинок в i-ій комірці, тобто число розміщень частинок дорівнює числу невід'ємних розв'язків рівняння $x_1 + x_2 + \cdots + x_m = n$. Обидва твердження випливають з тверджень теорем 1.11 та 1.12 відповідно. □

Модель розміщення частинок, яку ми розглянули, називається **моделлю Бозе-Ейнштейна**.

Розглянемо тепер розміщення однакових частинок коли виконуються принцип заборони Паулі, яка твердить що у кожній комірці не може бути більше однієї частинки. У цьому випадку маємо такий результат.

Теорема 1.16. *Нехай n однакових частинок розміщені в m комірках, $m \geq n$, коли у кожній комірці не може бути більше однієї частинки. Тоді число всіх можливих розміщень частинок дорівнює*

$$C_m^n.$$

Модель розміщення однакових частинок коли виконуються принцип заборони Паулі називається **моделлю Фермі-Дірака**.

Зауважимо, що, наприклад, електрони, протони та нейтрони у статистичній фізиці підпорядковані ститистиці Фермі-Дірака, а фотони та рі-мезони підпорядковані ститистиці Бозе-Ейнштейна.

Для порівняння результатів стосовно вибору n предметів із m предметів та результатів розміщення n частинок в m комірках збережемо їх у вигляді таблиць.

Результати стосовно вибору n предметів із m предметів зібрані в таблиці 1.2.

Вибірка	Впорядкована	Не впорядкована
З поверненням	m^n	C_{m+n-1}^n
Без поверненням	A_m^n	C_m^n

Таблиця 1.2

У тому випадку коли $n = 2$ та $m = 3$ відповідні вибірки висвітлені в таблиці 1.3.

Вибірка	Впорядкована	Не впорядкована
З поверненням	$(1,1)\ (1,2)\ (1,3)$ $(2,1)\ (2,2)\ (2,3)$ $(3,1)\ (3,2)\ (3,3)$	$[1,1]\ [1,2]\ [1,3]$ $\emptyset\ [2,2]\ [2,3]$ $\emptyset\ \emptyset\ [3,3]$
Без поверненням	$\emptyset\ (1,2)\ (1,3)$ $(2,1)\ \emptyset\ (2,3)$ $(3,1)\ (3,2)\ \emptyset$	$\emptyset\ [1,2]\ [1,3]$ $\emptyset\ \emptyset\ [2,3]$ $\emptyset\ \emptyset\ \emptyset$

Таблиця 1.3

Результати щодо розміщення n частинок в m комірках зібрані в таблиці 1.4.

Розміщення	Різні частинки	Однакові частинки
Без заборони	m^n	C^n_{m+n-1}
Із забороною	A^n_m	C^n_m

Таблиця 1.4

У тому випадку коли $n = 2$ та $m = 3$ відповідні результати висвітлені в таблиці 1.5.

Частинки	Різні	Однакові
Без заборони	$\|bc\|0\|0\|;\ \|b\|c\|0\|;\ \|b\|0\|c\|$ $\|c\|b\|0\|;\ \|0\|bc\|0\|;\ \|0\|b\|c\|$ $\|c\|0\|b\|;\ \|0\|c\|b\|;\ \|0\|0\|bc\|$	$\|bb\|0\|0\|;\ \|b\|b\|0\|;\ \|b\|0\|b\|$ $\emptyset\ \|0\|bb\|0\|;\ \|0\|b\|b\|$ $\emptyset\ \emptyset\ \|0\|0\|bb\|$
Із забороною	$\emptyset\ \|b\|c\|0\|;\ \|b\|0\|c\|$ $\|c\|b\|0\|;\ \emptyset\ \|0\|b\|c\|$ $\|c\|0\|b\|;\ \|0\|c\|b\|\ \emptyset$	$\emptyset\ \|b\|b\|0\|;\ \|b\|0\|b\|$ $\emptyset\ \emptyset\ \|0\|b\|b\|$ $\emptyset\ \emptyset\ \emptyset$

Таблиця 1.5

1.2.9 Поліноміальна теорема

Розглянемо питання про те, як розкривати дужки при обчисленні виразу виду $(a_1 + a_2 + \cdots + a_m)^n$.

Теорема 1.17. (Поліноміальна теорема.) *Вираз*

$$(a_1 + a_2 + \cdots + a_m)^n$$

дорівнює сумі всіх можливих доданків виду

$$\frac{n!}{k_1!k_2!\dots k_m!} a_1^{k_1} a_2^{k_2} \dots a_m^{k_m},$$

де $k_1 + k_2 + \cdots + k_m = n$, *тобто*

$$(a_1+a_2+\cdots+a_m)^n = \sum_{\substack{k_1\geq 0,\dots,k_m\geq 0\\ k_1+\cdots+k_m=n}} \frac{n!}{k_1!k_2!\dots k_m!} a_1^{k_1} a_2^{k_2} \dots a_m^{k_m}. \quad (1.11)$$

Доведення. Перемножимо $a_1 + a_2 + \cdots + a_m$ послідовно m раз. Тоді одержимо m^n доданків виду $d_1 d_2 \dots d_n$, де кожен множник дорівнює або a_1, або a_2,..., або a_m. Позначимо через $B(k_1, k_2, \dots, k_m)$ сукупність усіх тих доданків, де a_1 зустрічається множником k_1 раз, a_2 зустрічається множником k_2 раз,..., a_m зустрічається множником k_m раз. Число таких доданків дорівнює $C_n(k_1, k_2, \dots, k_m)$ — числу способів представлення множини $\{1, 2, \dots, n\}$ у вигляді суми m множин так, щоб множина B_s мала k_s елементів ($k_s \geq 0$, $k_1 + k_2 + \cdots + k_m = n$); множина B_s — це множина тих i для яких $d_i = a_s$). Згідно з теоремою 1.11

$$C_n(k_1, k_2, \dots k_m) = \frac{n!}{k_1!k_2!\dots k_m!}.$$

Отже, має місце рівність (1.11). □

При $m = 2$ рівність (1.11) набуває вигляду

$$\sum_{\substack{k_1+k_2=n\\ k_1\geq 0, k_2\geq 0}} \frac{n!}{k_1!k_2!}\, a_1^{k_1} a_2^{k_2} = (a_1 + a_2)^n. \quad (1.12)$$

Поклавши $k_2 = k, k_1 = n - k$, будемо мати

$$(a_1 + a_2)^n = \sum_{k=0}^{n} C_n^k a_1^{n-k} a_2^k.$$

Отже, формула біному Ньютона є частковим випадком поліноміальної теореми. Поклавши в (1.11) $a_1 = a_2 = \dots a_m = 1$ отримаємо ще одне доведення рівності (1.12).

Приклад 1.26. Скільки доданків в правій частині розкладу (1.11)?

Число доданків дорівнює числу невід'ємних розв'язків рівняння $k_1 + k_2 + \ldots k_m = n$, тобто дорівнює

$$f_m^n = C_{n+m-1}^{m-1} = C_{n+m-1}^{n}.$$

1.3 Аксіоми теорії ймовірностей. Властивості ймовірності

1.3.1 Аксіоми класу випадкових подій

Нехай $\mathfrak{F}$ – клас усіх випадкових подій у деякому стохастичному експерименті із простором елементарних подій Ω.

Виділимо такі властивості випадкових подій:

1 . (F_1 – нормованість) Вся множина $\Omega \in \mathfrak{F}$.

2. (F_2 – доповнення) З того що $A \in \mathfrak{F}$ випливає, що $\overline{A} \in \mathfrak{F}$.

3. (F_3– об'єднання) З того що $A, B \in \mathfrak{F}$ випливає, що $A \cup B \in \mathfrak{F}$.

4. (F_4 – зліченні об'єднання) Якщо $A_n \in \mathfrak{F}$ – зліченна послідовність, то $\cup_n A_n \in \mathfrak{F}$.

Клас підмножин універсальної множини Ω, який задовольняє умови (F_1) – (F_3), називається алгеброю підмножин Ω.

Клас підмножин універсальної множини Ω, який задовольняє умови (F_1) – (F_4), називається сигма-алгеброю (чи σ-алгеброю) підмножин Ω.

Властивість (F_3) еквівалентна умові

5. (F_5 – скінченні об'єднання) З того що $A_k \in \mathfrak{F}, k = 1, \ldots, n$, випливає, що $\cup_{k=1}^n A_k \cup B \in \mathfrak{F}$.

Приклад 1.27. Для довільної події $A \in \Omega$ клас $\{\varnothing, A, \overline{A}, \Omega\}$ є алгеброю і одночасно σ-алгеброю.

Першу групу аксіом теорії ймовірностей можна сформулювати так: "Клас $\mathfrak{F}$ усіх випадкових подій є сигма-алгеброю".

Для побудови сигма-алгебр, виходячи із заданих алгебр, використовують такий метод.

Означення 1.6. Сигма-алгебра $\mathfrak{F}$ породжена класом випадкових подій $\mathfrak{U}$, позначається як $\mathfrak{F} = \sigma[\mathfrak{U}]$, якщо вона є найменшою з тих сигма-алгебр, що містять цей клас, тобто $\mathfrak{F}$ є перетином всіх сигма-алгебр, що містять цей клас.

Зауваження 1.1. З наведених аксіом випливають всі вказані вище властивості класу випадкових подій.

Зауваження 1.2. Визначення породженої сигма-алгебри коректне, оскільки хоча б одна вказана сигма-алгебра S існує (наприклад, клас усіх підмножин Ω), а переріз будь-якої кількості сигма-алгебр є сигма-алгеброю.

Зауваження 1.3. Аксіоми (F_1) – (F_3) алгебри підмножин Ω можна замінити на такі: (а) Вся множина $\Omega \in \mathfrak{F}$; (б) З тоо що $A, B \in \mathfrak{F}$ випливає, що $A \setminus B \in \mathfrak{F}$.

Вправа 1.2. Нехай $\Omega = \{1, 2, 3, 4, 5, 6\}$. Знайти всі елементи найменшої алгебри, яка містить події $A = \{1, 2\}$, $B = \{, 4, 5, 6\}$.

1.3.2 Аксіоми ймовірності

У теорії ймовірностей із кожною випадковою подією пов'язують числову міру її вірогідності – ймовірність.

Означення 1.7. Імовірністю називається числова функція $\mathbb{P}$ на класі $\mathfrak{F}$ всіх випадкових подій з такими властивостями:

(P1 – невід'ємність). Тобто $\mathbb{P}(A) \geq 0, \forall A \in \mathfrak{F}$.

(P2 – нормованість). Тобто $\mathbb{P}(\Omega) = 1$.

(P3 – сигма-адитивність). Для послідовності $\{A_n, n \geq 1\}$ попарно несумісних випадкових подій (тобто $A_i \cap A_j = \varnothing, \forall i \neq j$) справедлива рівність

$$\mathbb{P}(\cup_{n \geq 1} A_n) = \sum_{n \geq 1} \mathbb{P}(A_n).$$

Зауваження 1.4. Зауважимо, що твердження аксіоми (P3) еквівалентне такій парі тверджень:

(P3' – адитивність) Для несумісних подій A, B (тобто $A \cap B = \varnothing$) справедлива рівність

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B),$$

(P4 – неперервність в нулі). Для довільної послідовності подій $\{A_n, n \geq 1\}$ таких, що $A_{n+1} \subset A_n$, та $\cap_{n \geq 1} A_n = \varnothing$, має місце збіжність

$$\lim_{n \to \infty} \mathbb{P}(A_n) = 0.$$

Означення 1.8. Функція P називається адитивною ймовірністю, якщо замість аксіоми (P3) виконується слабша умова (P3').

За індукцією неважко довести (вправа), що властивість адитивності (Р3') еквівалентна скінченній адитивності:

(Р3" – скінченна адитивність) Для довільної скінченної послідовності попарно несумісних подій $\{A_k,\, 1 \le k \le n\}$

$$\mathbb{P}(\cup_{1\le k\le n} A_k) = \sum_{1\le k\le n} \mathbb{P}(A_k).$$

Отже другу групу аксіом теорії ймовірностей можна сформулювати так:

"Імовірність є невід'ємною нормованою сигма-адитивною функцією на класі всіх випадкових подій".

Аксіоматичне означення теорії ймовірностей означає існування ймовірнісного простору

$$(\Omega, \mathfrak{F}, \mathbb{P})\,,$$

що має такі елементи:

Ω– простір елементарних подій (абстрактна множина),

$\mathfrak{F}$ – клас усіх випадкових подій – підмножин простору Ω, які утворюють сигма-алгебру з властивостями (F1 – F3), та

$\mathbb{P}$ – імовірність на сигма-алгебрі подій $\mathfrak{F}$ із властивостями (Р1 – Р3).

1.3.3 Властивості ймовірності

Властивості ймовірності

З аксіом теорії ймовірностей випливає, що справедливі такі властивості ймовірності.

1. Імовірність неможливої події: $\mathbb{P}(\varnothing) = 0$. Справді,

$$\Omega + \varnothing = \Omega, \quad \Omega \cap \varnothing = \varnothing,$$

події Ω і $\varnothing$ несумісні, $\Omega \cap \varnothing = \varnothing$, тому

$$1 = \mathbb{P}\left(\Omega\right) + \mathbb{P}\left(\varnothing\right) = \mathbb{P}\left(\Omega\right) = 1.$$

Із цієї рівності випливає, що

$$\mathbb{P}\left(\varnothing\right) = 0.$$

Зауважимо також, що з $\mathbb{P}\left(A\right) = 0$ не випливає, що A – неможлива подія.

2. Імовірність доповнення події $A \in \mathfrak{F}$

$$\mathbb{P}\left(\overline{A}\right) = 1 - \mathbb{P}(A),$$

оскільки $\Omega = A \cup \overline{A}$, де доданки несумісні, $A \cap \bar{A} = \varnothing$, отже

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(A \cup \overline{A}\right) = \mathbb{P}\left(\overline{A}\right) + \mathbb{P}(A).$$

3. Імовірність вкладеної різниці: для $A \subset B$, $A, B \in \mathfrak{F}$

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A),$$

оскільки $B = A \cup (B \setminus A)$, де доданки несумісні:

$$\mathbb{P}(B) = \mathbb{P}(A \cup (B \setminus A)) = \mathbb{P}(A) + \mathbb{P}(B \setminus A).$$

4. Монотонність імовірності: для $A \subset B$, $A, B \in \mathfrak{F}$

$$\mathbb{P}(A) \leq \mathbb{P}(B),$$

виводиться з невід'ємності та попереднього пункту:

$$\mathbb{P}(B) - \mathbb{P}(A) = \mathbb{P}(B \setminus A) \geq 0.$$

5. Множина значень імовірності: для $A \in \mathfrak{F}$

$$0 \leq \mathbb{P}(A) \leq 1,$$

оскільки $\varnothing \subset A \subset \Omega$ і має місце монотонність імовірності.

6. Імовірність об'єднання двох подій $A, B \in \mathfrak{F}$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B),$$

оскільки $A \cup B = A \cup (B \setminus (A \cap B))$, де доданки несумісні, а ймовірність другого визначається властивістю ймовірності вкладеної різниці:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (B \setminus (A \cap B))) =$$

$$= \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

7. Формула включення–виключення для подій $\{A_k, 1 \leq k \leq n\}$:

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k) - \sum_{1 \leq i \leq j \leq n} \mathbb{P}(A_i \cap A_j) +$$

$$+\sum_{1\le i\le j\le k\le n}\mathbb{P}(A_i\cap A_j\cap A_k)-\cdots+(-1)^{n-1}\mathbb{P}(A_1\cap\cdots\cap A_n).$$

При $n=2$ це є властивість 6, а для $n>2$ формула доводиться за індукцією з властивості ймовірності об'єднання.

8. Напівадитивність імовірності:

$$\mathbb{P}\left(\bigcup_{k=1}^{n}A_k\right)\le\sum_{k=1}^{n}\mathbb{P}(A_k).$$

Доведення. Для доведення позначимо

$$B_1=A_1,\quad B_k=A_k\setminus(\cup_{j=1}^{k-1}A_j).$$

Тоді події B_k попарно несумісні, $B_k\subset A_k$ і

$$\cup_{k=1}^{n}A_k=\cup_{k=1}^{n}B_k.$$

Звідси за адитивністю та монотонністю ймовірностей

$$\mathbb{P}(\cup_{k=1}^{n}A_k)=\mathbb{P}(\cup_{k=1}^{n}B_k)=$$

$$=\sum_{k=1}^{n}\mathbb{P}(B_k)\le\sum_{k=1}^{n}\mathbb{P}(A_k).$$

□

9. Неперервність імовірності.
(а) Якщо $A_n\uparrow A$, то $\mathbb{P}(A_n)\uparrow\mathbb{P}(A)$.
(б) Якщо $A_n\downarrow A$, то $\mathbb{P}(A_n)\downarrow\mathbb{P}(A)$.

Доведення. Для доведення (а) позначимо $B_1=A_1$, $B_n=A_n\setminus A_{n-1}, n>1$. Тоді B_n попарно несумісні,

$$\cup_{k\le n}B_k=A_n,\quad \cup_k B_k=A,$$

$$\mathbb{P}(A)=\sum_n\mathbb{P}(B_n)=\lim_{n\to\infty}\sum_{k\le n}\mathbb{P}(B_k)=\lim_{n\to\infty}\mathbb{P}(A_n),$$

де остання рівність виводиться з адитивності, а збіжність є монотонною.

Для доведення (б) застосуємо (а) для доповнень: $\overline{A_n}\uparrow\overline{A}$,

$$1-\mathbb{P}(A_n)=\mathbb{P}(\overline{A_n})\uparrow\mathbb{P}(\overline{A})=1-\mathbb{P}(A).$$

□

10. Зліченна напівадитивність імовірності:

$$\mathbb{P}(\cup_k A_k) \le \sum_k \mathbb{P}(A_k)$$

випливає із напівадитивності та неперервності ймовірності:

$$\mathbb{P}(\cup_k A_k) = \lim_{n\to\infty} \mathbb{P}(\cup_{1\le k\le n} A_k) \le \lim_{n\to\infty} \sum_{1\le k\le n} \mathbb{P}(A_k).$$

11. Еквівалентність неперервності в нулі та сигма-адитивності для адитивної ймовірності.

Для того, щоб адитивна ймовірність $\mathbb{P}$ на алгебрі подій $\mathfrak{U}$ була сигма-адитивною, необхідно і достатньо, щоб вона була неперервною в нулі на $\mathfrak{U}$, тобто для будь-якої послідовності подій $A_n \in \mathfrak{U}$ таких, що $A_n \downarrow \varnothing$, повинна мати місце збіжність $\mathbb{P}(A_n) \downarrow 0$.

Доведення. Якщо функція $\mathbb{P}$ сигма-адитивна, то вона неперервна в нулі за властивістю неперервності ймовірності.

Нехай функція $\mathbb{P}$ адитивна на алгебрі $\mathfrak{U}$ та неперервна в нулі, а $A_n \in \mathfrak{U}$ попарно несумісні події. Позначимо $B_n = \cup_{k>n} A_k$. Тоді $B_n \downarrow \varnothing$, та $\mathbb{P}(B_n) \downarrow 0$. Оскільки внаслідок скінченної адитивності

$$\mathbb{P}(\cup_k A_k) = \mathbb{P}(\cup_{k\le n} A_k) + \mathbb{P}(B_n) = \sum_{k\le n} \mathbb{P}(A_k) + \mathbb{P}(B_n) \to \sum_{k\ge 1} \mathbb{P}(A_k),$$

при $n \to \infty$, то ймовірність $\mathbb{P}$ сигма-адитивна на алгебрі $\mathfrak{U}$. □

12. Продовження неперервної ймовірності на сигма-алгебру.

Для того, щоб адитивна ймовірність $\mathbb{P}$ на алгебрі подій $\mathfrak{U}$ мала продовження до сигма-адитивної ймовірності на породженій сигма-алгебрі $\mathfrak{F} = \sigma[\mathfrak{U}]$, необхідно і достатньо, щоб вона була неперервною в нулі на алгебрі $\mathfrak{U}$.

Доведення. Якщо ймовірність $\mathbb{P}$ сигма-адитивна, то вона неперервна в нулі за властивістю неперервності. Нехай ймовірність $\mathbb{P}$ адитивна на алгебрі $\mathfrak{U}$ та неперервна в нулі. Тоді $\mathbb{P}$ сигма-адитивна на алгебрі $\mathfrak{U}$, як показано у попередній властивості. За цієї умови існування та єдиність сигма-адитивного продовження на σ -алгебру $\mathfrak{F} = \sigma[\mathfrak{U}]$ стверджується у теоремі Каратеодорі про продовження міри, яка доводиться в курсі теорії міри та інтеграла. □

Вправа 1.3. Довести: (а) рівність

$$\mathbb{P}(A \triangle B) = \mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A \cap B),$$

(б) нерівність трикутника

$$\mathbb{P}(A \triangle B) \le \mathbb{P}(A \triangle C) + \mathbb{P}(C \triangle B),$$

(в) нерівність різниць

$$\mathbb{P}((\cup A_k) \triangle (\cup B_k)) \le \sum_k \mathbb{P}(A_k \triangle B_k).$$

1.4 Приклади ймовірнісних просторів

Розглянемо конкретні приклади ймовірнісних просторів.

1.4.1 Класичне означення ймовірностей

Нехай простір $\Omega = \{\omega_1, \ldots, \omega_n\}$ – скінченний, а клас подій $\mathfrak{F}$ містить всі підмножини Ω.

Припустимо, що всі елементарні події рівноймовірні (рівноможливі, симетричні, статистично нерозрізненні). Тоді їх імовірності мають бути однаковими і в сумі повинні дорівнювати 1. Отже,

$$\mathbb{P}(\{\omega_k\}) = \frac{1}{n}, \quad k = 1, \ldots, n,$$

то ймовірність будь-якої події має дорівнювати сумі доданків $1/n$ по всіх m елементарних подіях, що їй сприяють, тобто дорівнювати відношенню m/n. Звідси дістанемо таке означення.

Означення 1.9. Будемо говорити, що для даного ймовірнісного простору $(\Omega, \mathfrak{F}, \mathbb{P})$ зі скінченним простором елементарних подій $\Omega = \{\omega_1, \ldots, \omega_n\}$ прийняте класичне означення ймовірностей, якщо для будь-якої події $A \subset \Omega$ її ймовірність дорівнює відношенню кількості елементарних подій, що сприяють A, до кількості усіх елементарних подій:

$$\mathbb{P}(A) = \frac{N(A)}{N(\Omega)},$$

де $N(A)$ – кількість елементів (потужність) множини A.

Зауваження 1.5. Якщо у формулюванні задачі явно не вказано на вибір означення ймовірностей, однак $N(\Omega) < \infty$ і присутні ключові слова випадково, навмання ... – то треба обрати саме класичне означення.

Приклад 1.28. Підкидання двох монет. Простір всіх елементарних подій $\Omega = \{AA, AP, PA, PP\}$ складається із 4 елементарних подій, $N(\Omega) = 4$. Події B {при підкиданні двох монет першим випаде A} відповідає підмножина $B = \{AA, AP\}$ простору Ω із 2 елементарних подій, $N(B) = 2$. Події C {принаймі один раз випаде A} відповідає підмножина $C = \{AA, AP, PA\}$ простору Ω із 3 елементарних подій, $N(C) = 3$. Тому

$$\mathbb{P}(B) = \frac{2}{4}, \quad \mathbb{P}(C) = \frac{3}{4}.$$

Приклад 1.29. Підкидання грального кубика. Простір всіх елементарних подій $\Omega = \{1, 2, 3, 4, 5, 6\}$. Події B {при підкиданні грального кубика випаде парна кількість очок} відповідає підмножина $B = \{2, 4, 6\}$ простору Ω із 3 елементарних подій, $N(B) = 3$. Події C {при підкиданні грального кубика випаде непарна кількість очок} відповідає підмножина $C = \{1, 3, 5\}$ простору Ω із 3 елементарних подій, $N(C) = 3$.

$$\mathbb{P}(B) = \frac{3}{6} = \frac{1}{2}, \quad \mathbb{P}(C) = \frac{3}{6} = \frac{1}{2}.$$

Приклад 1.30. Розглянемо тепер результати кидання двох кубиків. Якщо кубики правильні, то випадання кожної із 36 можливих комбінацій числа очок на першому і на другому кубиках можна вважати рівноможливим. Наприклад, імовірність випадання 12 очок дорівнює $\frac{1}{36}$. Випадання 11 очок, тобто однієї шістки й однієї п'ятірки, можливе двома способами: шістка на першому кубику і п'ятірка на другому, і навпаки, п'ятірка на першому і шістка на другому. Тому ймовірність випадання одинадцяти очок дорівнює $\frac{2}{36} = \frac{1}{18}$. Ймовірності випадання тієї або іншої суми очок подаються такою таблицею:

К-ть очок	2	3	4	5	6	7	8	9	10	11	12
Імовірність	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Таблиця 1.6

Приклад 1.31. Із колоди карт (36 карт) наугад вибирають 3 карти. Знайти ймовірність того, що серед них буде рівно один туз.

Розв'язування. Повна група рівноможливих і несумісних подій в нашій задачі складається з усіх можливих комбінацій по 3 карти, їхнє число дорівнює C_{36}^3. Число сприятливих випадків можна підрахувати так. Одного туза можна вибрати C_4^1 різними способами, а дві інші карти (не тузи) можна вибрати C_{32}^2 різними способами. Оскільки для кожного визначеного туза дві інші карти можуть бути вибрані C_{32}^2 різними способами, то всіх сприятливих випадків буде $C_4^1 \cdot C_{32}^2$. Отже, шукана ймовірність дорівнює:

$$P = \frac{C_4^1 \cdot C_{32}^2}{C_{36}^3} = \frac{4 \cdot \frac{32 \cdot 31}{2 \cdot 1}}{\frac{36 \cdot 35 \cdot 34}{3 \cdot 2 \cdot 1}} = \frac{31 \cdot 16}{35 \cdot 3 \cdot 17} = \frac{496}{1785} \approx 0,27781,$$

тобто трохи більша за 0,25.

Приклад 1.32. Із колоди карт (36 карт) наугад виймають три карти. Знайти ймовірність того, що серед них буде хоча б один туз.

Розв'язування. Позначимо подію, яка нас цікавить, буквою A, вона може бути зображена як сума трьох таких несумісних подій: A_1 – поява одного туза, A_2 – поява двох тузів, A_3 – поява трьох тузів. Міркуваннями, аналогічними тим, які були проведені при розв'язуванні попередньої задачі, легко встановити, що число випадків,

сприятливих події A_1, дорівнює $C_4^1 \cdot C_{32}^2$,
сприятливих події A_2, дорівнює $C_4^2 \cdot C_{32}^1$,
сприятливих події A_3, дорівнює $C_4^3 \cdot C_{32}^0$.

Оскільки число всіх можливих випадків дорівнює C_{36}^3, то

$$P(A_1) = \frac{C_4^1 \cdot C_{32}^2}{C_{36}^3} = \frac{16 \cdot 31}{3 \cdot 35 \cdot 17} \approx 0,2778,$$

$$P(A_2) = \frac{C_4^2 \cdot C_{32}^1}{C_{36}^3} = \frac{3 \cdot 16}{3 \cdot 35 \cdot 17} \approx 0,0269,$$

$$P(A_3) = \frac{C_4^3 \cdot C_{32}^0}{C_{36}^3} = \frac{1}{3 \cdot 35 \cdot 17} \approx 0,0006.$$

На підставі теореми додавання

$$P(A) = P(A_1) + P(A_2) + P(A_3) = \frac{109}{3 \cdot 119} \approx 0,3053.$$

Цей приклад можна розв'язати й іншим методом. Подія $\bar{A}$, протилежна події A, полягає в тому, що серед вибраних карт не буде жодного туза. Очевидно, що три нетузи можна вийняти з колоди карт C_{32}^3 різними способами, і отже,

$$P\left(\bar{A}\right) = \frac{C_{32}^3}{C_{36}^3} = \frac{32 \cdot 31 \cdot 30}{36 \cdot 35 \cdot 34} = \frac{31 \cdot 8}{3 \cdot 17 \cdot 7} \approx 0,6947.$$

Шукана ймовірність дорівнює

$$P\left(A\right) = 1 - P\left(\bar{A}\right) = 1 - \frac{C_{32}^3}{C_{36}^3} \approx 0,3053.$$

Зауваження 1.6. В обох прикладах вираз "наугад" означав, що всі можливі комбінації по три карти рівноймовірні.

Приклад 1.33. Колоду, що складається із 36 карт, наугад, поділяють на дві рівні частини. Чому дорівнює ймовірність того, що в обох частинах буде по рівному числу червоних і чорних карт?

Розв'язування. Нам треба знайти ймовірність того, що серед наугад вибраних із колоди 18 карт 9 будуть червоними і 9 чорними. Загальне число різних способів, якими можна вибрати 18 карт із 36, дорівнює C_{36}^{18}. Сприятливими способами будуть усі ті, при яких із 18 червоних карт будуть вибрані 9 та із 18 чорних карт будуть вибрані 9.

C_{18}^9 різними способами можна вибрати 9 червоних карт із 18 червоних карт. Також C_{18}^9 різними способами можна вибрати 9 чорних карт із 18 чорних карт. Оскільки при вийманні 9 визначених червоних карт 9 чорних можна витягти C_{18}^9 різними способами, загальне число сприятливих способів дорівнює $C_{18}^9 \cdot C_{18}^9$. Шукана ймовірність, отже, дорівнює

$$p = \frac{C_{18}^9 \cdot C_{18}^9}{C_{36}^{18}} = \frac{(18!)^4}{36!(9!)^4}.$$

Щоб уявити собі, яке значення має ця ймовірність і при цьому не робити складних обчислень, скористаємося формулою Стірлінга, згідно з якою має місце така асимптотична рівність:

$$n! \approx \sqrt{2\pi n} n^n e^{-n}.$$

Отже, маємо

$$18! \approx 18^{18} e^{-18} \sqrt{2\pi \cdot 18},$$

$$9! \approx 9^9 e^{-9}\sqrt{2\pi \cdot 9},$$
$$36! \approx 36^{36} e^{-36}\sqrt{2\pi \cdot 36},$$

і, отже,

$$p \approx \frac{\left(\sqrt{2\pi \cdot 18} \cdot 18^{18} e^{-18}\right)^4}{\sqrt{2\pi \cdot 36} \cdot 36^{36} \cdot e^{-36}\left(\sqrt{2\pi \cdot 9} \cdot 9^9 \cdot e^{-9}\right)^4}.$$

Після нескладних перетворень знаходимо, що

$$p \approx \frac{2}{\sqrt{18\pi}} \approx \frac{4}{15} \approx 0,26.$$

Приклад 1.34. Маємо n частинок, кожна з яких може бути з однією і тією ж імовірністю $\frac{1}{N}$ у кожній з N, $N > n$, комірок. Знайти ймовірність того, що:

1) у визначених n комірках буде по одній частинці,
2) в яких-небудь n комірках буде по одній частинці.

Розв'язування. Ця задача відіграє важливу роль у сучасній статистичній фізиці і залежно від того, як утворюється повна група рівноможливих подій, приходять до тієї або іншої фізичної статистики: Больцмана, Бозе – Ейнштейна, Фермі – Дірака.

У статистиці Больцмана рівноможливими вважаються розміщення, що відрізняються не лише кількістю, а й індивідуальністю частинок; у кожній комірці може міститися довільне число частинок від 0 до n. Загальне число можливих розміщень підрахуємо так: кожна частинка може міститися в одній з N комірок, отже, n частинок можна розмістити по комірках N^n різними способами.

У першому випадку кількість сприятливих подій буде, очевидно, $n!$ і, отже, імовірність того, що у визначені n комірок потрапить по одній частинці, дорівнює

$$p_1 = \frac{n!}{N^n}.$$

У другому випадку число сприятливих подій буде в C_N^n разів більше, отже, імовірність того, що в будь-які п комірок потрапить по одній частинці, дорівнює

$$p_2 = \frac{n! C_N^n}{N^n} = \frac{N!}{N^n\,(N-n)!}.$$

У статистиці Бозе – Ейнштейна вважаються тотожними випадки, коли частинки міняються місцями (важливо лише, скільки частинок потрапило в комірку, а не індивідуальність частинок), і повна група подій складається з усіх можливих розміщень n частинок в N комірках, причому за одне розміщення вважається цілий клас больцманівських розміщень, що відрізняються не кількістю частинок, які містяться в певних комірках, а лише самими частинками.

Для наочного уявлення про різницю між статистиками Больцмана і Бозе – Ейнштейна розглянемо випадок: $N = 4, n = 2$. Усі можливі в цьому випадку розміщення можна записати у вигляді таблиці 1.7, в якій a і b – назви частинок. У статистиці Больцмана всі 16 можливостей являють собою різні рівноможливі події, у статистиці Бозе – Ейнштейна випадки 5 і 11, 6 і 12, 7 і 13, 8 і 14, 9 і 15, 10 і 16 попарно ототожнюються, отже, маємо групу із 10 рівноможливих подій.

1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	ab				a	a	a				b	b	b			
2		ab			b			a	a		a			b	b	
3			ab			b		b		a		a		a		b
4				ab			b		b	b			a		a	a

Таблиця 1.7

Обчислимо тепер загальне число рівноймовірних випадків у статистиці Бозе – Ейнштейна. Для цього зауважимо, що всі можливі розміщення частинок по комірках можна знайти так: розмістимо комірки на прямій щільно одна до одної, а далі поряд одну біля одної на тій самій прямій розмістимо наші частинки. Розглянемо тепер можливі перестановки частинок і перегородок між комірками. Легко зрозуміти, будуть враховані всі можливі заповнення комірок, що відрізняються як порядком розміщення частинок у комірках, так і порядком розміщення перегородок.

Число цих перестановок дорівнює $(N - n + 1)!$. Серед цих перестановок є й тотожні: кожне розміщення по комірках лічимо $(N-1)!$ разів, оскільки ми відрізняли, які перегородки були між комірками, а, крім того, кожне розміщення по комірках знову лічимо по $n!$ разів, оскільки було враховано не лише кількість частинок у комірці, а й те, які це частинки і в якому порядку вони розміщені. Отже, кожне розміщення по комірках ми лічили $n!(N-1)!$ разів, звідси число різних у розумінні Бозе – Ейнштейна розміщень по комірках дорівнює

$$\frac{(n+N-1)!}{n!\,(N-1)!}.$$

Таким чином, кількість рівноймовірних подій в повній групі подій нами знайдена. Тепер легко можна відповісти на запитання нашої задачі. У статистиці Бозе – Ейнштейна ймовірності p_1 і p_2 дорівнюють

$$p_1 = \frac{1}{\frac{(n+N-1)!}{n!(N-1)!}} = \frac{n!\,(N-1)!}{(n+N-1)!},$$

$$p_2 = \frac{C_N^n}{\frac{(n+N-1)!}{n!(N-1)!}} = \frac{N!\,(N-1)!}{(N-n)!\,(n+N-1)!}.$$

Розглянемо, нарешті, статистику Фермі – Дірака. Згідно з цією статистикою в комірці може бути або одна, або не бути жодної частинки. Індивідуальність частинок не береться до уваги. Загальну кількість різних розміщень частинок по комірках у статистиці Фермі – Дірака обчислити легко: перша частинка може бути розміщена N різними способами, друга – лише $N-1$, третя – $(N-2)$ і, нарешті, остання – $(N-n+1)$ різними способами. Таким чином, n частинок по N комірках можуть бути розміщені

$$N\,(N-1)\cdots(N-n+1) = \frac{N!}{(N-n)!}$$

способами. Але при цьому за різні способи тут не вважаються ті, які відрізняються лише перестановкою частинок. Тому рівноможливих випадків маємо всього

$$\frac{N!}{n!\,(N-n)!} = C_N^n.$$

Легко переконатися, що в статистиці Фермі – Дірака шукані ймовірності дорівнюють

$$p_1 = \frac{n!\,(N-n)!}{N!},\ p_2 = 1.$$

Розглянутий приклад показує, наскільки важливо точно визначити, які події вважаються рівноможливими. Також ми бачимо, що саме поняття рівноймовірності подій залежно від фізичних умов задачі може набувати різного змісту.

1.4.2 Дискретний імовірнісний простір

Не завжди можна вважати, що елементарні події рівноможливі. В деяких випадках можна застосувати таке узагальнення.

Нехай простір $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$ – дискретний (скінченний або зліченний), клас випадкових подій $\mathfrak{F}$ містить всі підмножини Ω, і задана певна ймовірність $\mathbb{P}$ на $\mathfrak{F}$. Тоді визначені ймовірності окремих елементарних подій $p_k = \mathbb{P}(\{\omega_k\})$.

Оскільки для довільної зліченної множини $A = \cup_{k:\omega_k \in A}\{\omega_k\}$, то за властивістю сигма-адитивності

$$\mathbb{P}(A) = \sum_{k:\omega_k \in A} p_k, \quad p_k = \mathbb{P}(\{\omega_k\}).$$

Числова послідовність $p_k, k \geq 1$, задовольняє умови

$$p_k \geq 0, \quad \sum_{k=1}^{\infty} p_k = 1$$

і називається дискретним розподілом імовірностей.

Отже, для кожної ймовірності $\mathbb{P}$ на $\mathfrak{F}$ визначений дискретний розподіл на одноточкових множинах.

Навпаки, для кожного дискретного розподілу $\{p_k, k \geq 1\}$ наведена формула для ймовірності $\mathbb{P}(A)$ задає загальне означення ймовірностей у дискретному ймовірнісному просторі.

Теорема 1.18. (Теорема про сигма-адитивність дискретної ймовірнісної міри.) *Для кожного дискретного розподілу ймовірностей $\{p_k, k \geq 1\}$ функція*

$$\mathbb{P}(A) = \sum_{k:\omega_k \in A} p_k, \quad A \in \mathfrak{F}$$

є адитивною ймовірністю та неперервною в нулі, а тому є ймовірністю.

Доведення. Доведення. Невід'ємність та нормованість очевидні. Адитивність є наслідком адитивності суми. Еквівалентність неперервності в нулі та сигма-адитивності встановлено у теоремі про основні властивості ймовірностей □

Приклад 1.35. Підкидання монети до першого успіху. Простір елементарних подій має вигляд $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$. Співставимо елементарну подію ω_n з елементарною подією в серії з n підкидань монети, що містить $n-1$ реверс (неуспіх) та 1 аверс (успіх). Виходячи з цього співставлення, природно постулювати, що

$$p_n = \mathbb{P}(\{\omega_n\}) = 2^{-n},\ n \geq 1.$$

Ця послідовність дійсно є дискретним розподілом. Тому ймовірність парної кількості підкидань до першого успіху дорівнює

$$\sum_{n=1}^{\infty} 2^{-2n} = \frac{1}{3}.$$

А ймовірність непарної кількості підкидань до першого успіху дорівнює

$$\sum_{n=1}^{\infty} 2^{-2n+1} = \frac{2}{3}.$$

1.4.3 Геометричні ймовірності

На прикладі задачі про випадковий вибір точки на відрізку $[a, b)$ бачимо, що існують стохастичні експерименти з незліченною кількістю елементарних подій. Хоча окремі елементарні події тут виглядають рівноможливими, класичне означення застосувати неможливо – всі ймовірності елементарних подій мають бути нульовими, тому що сума незліченної кількості додатних чисел нескінченна. Тому природно перейти від імовірностей окремих точок до ймовірностей інтервалів.

Нехай $\Omega = [a, b)$, сигма-алгебра подій $\mathfrak{F}$ містить всі інтервали $[x, y) \subset [a, b)$, і на $\mathfrak{F}$ задана деяка ймовірність $\mathbb{P}$. Припущення про рівноможливість окремих точок можна проінтерпретувати так, що ймовірність потрапляння випадкової точки до певного інтервалу залежить лише від його довжини, а не від розташування цього інтервалу.

З цього припущення, адитивності, нормованості та неперервності ймовірності випливає, що

$$\mathbb{P}([x, y)) = \frac{y-x}{b-a} = \frac{l[x, y)}{l[a, b)}.$$

У загальному випадку за припущення рівноможливості приходимо до такого означення.

Означення 1.10. Нехай простір $\Omega \subset \mathbb{R}^d$ є обмеженою множиною евклідового простору $\mathbb{R}^d$, а клас $\mathfrak{F}$ містить усі підмножини Ω із визначеним d-вимірним об'ємом. Будемо говорити, що у просторі $(\Omega, \mathfrak{F}, \mathbb{P})$ прийнято геометричне означення ймовірностей, якщо для всіх $A \in \mathfrak{F}$

$$\mathbb{P}(A) = \frac{\mathrm{vol}(A)}{\mathrm{vol}(\Omega)},$$

де $\mathrm{vol}(A)$ – d-вимірний об'єм множини A, тобто довжина при $d = 1$, площа при $d = 2$, об'єм при $d = 3$ і так далі.

Зауваження 1.7. Якщо у формулюванні задачі явно не вказано на вибір означення ймовірностей, але $\Omega \subset \mathbb{R}^d$ і присутні ключові слова випадково, навмання – то треба обрати саме геометричне означення.

Приклад 1.36. Задача про зустріч. Дві особи A і B домовилися зустрітися в певному місці між 12-ю годиною і 1-ю. Той, хто приходить першим, чекає на другого протягом 20 хв. і якщо не дочекається, іде. Чому дорівнює ймовірність зустрічі осіб A і B, якщо прихід кожної з них протягом указаної години може відбутися навгад і моменти приходу незалежні.

Розв'язування. Позначимо момент приходу особи A через x, а момент приходу особи B через y. Множина Ω складається із точок квадрата

$$\Omega = \{(x, y) \colon 0 \le x \le 60, 0 \le y \le 60\}.$$

Множина $\mathfrak{F}$ складається з усіх вимірних (борелівських) множин, утворених із точок квадрата. Зокрема, входить до складу $\mathfrak{F}$ і є випадковою подією множина, що складається із точок замкнутої області

$$C = \{(x, y) \colon 0 \le x \le 60, 0 \le y \le 60, |x - y| \le 20\}.$$

Зважаючи на незалежність моментів приходу осіб A і B, ймовірність того, що особа A прийде в проміжок від x до $x + h$, а особа B – у проміжок від y до $y + s$, дорівнює

$$\frac{h}{60} \cdot \frac{s}{60},$$

тобто пропорційна площі прямокутника із сторонами h і s. Звідси виводимо, що шукана ймовірність дорівнює відношенню площі області C до площі всього квадрата, тобто

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

Вправа 1.4. Яка ймовірність того, що з трьох наудачу узятих відрізків довжини не більше a можна побудувати трикутник?

1.4.4 Статистичне (частотне) означення ймовірності

Тривалі спостереження над появою або непоявою події A при великій кількості повторних випробовувань, що відбуваються при незмінному комплексі умов Ξ, показують, що число появ події A підкоряється стійким закономірностям. А саме, якщо через μ позначимо число появ події A при n незалежних випробовуваннях, то з'ясовується, що відношення $\frac{\mu}{n}$ (частота події A) при досить великих n для більшості таких серій спостережень зберігає майже сталу величину, причому більші відхилення спостерігаються тим рідше, чим численнішими є проведені випробовування. Цей емпіричний факт має глибокі основи в теоремі Бернуллі. Те, що при великій кількості випробовувань частота події залишається майже незмінною, дає нам змогу розширити коло явищ, про ймовірність яких будемо говорити.

На закінчення зупинимося на досить поширеній, особливо серед природознавців, концепції ймовірності, яку дав Р. Мізес. Згідно з Мізесом, якщо частота зі збільшенням кількості досліджень дедалі менше відхиляється від імовірності p, то має виконуватися співвідношення

$$p = \lim_{n\to\infty} \frac{\mu}{n}.$$

Не будемо зупинятися на деталях теорії Р. Мізеса, зокрема на тих обмеженнях і умовах, які вона додатково накладає на послідовність випробовувань. За деталями його теорії відішлемо читача до книги Р. Мизеса "Вероятность и статистика" Изд. 5-е. — М.: Либроком, 2009. — 264 с., а також до статті О.Я. Хінчина // Успехи физических наук. – 1929. – Вип. 2. – Т. IX. – С. 141–166.

1.5 Умовна ймовірність та найпростіші основні формули

Уже йшлося про те, що в основі означення ймовірності події лежить певна сукупність умов Ξ. Якщо ніяких обмежень при обчисленні ймовірності $P(A)$ не накладається, то такі ймовірності називаються безумовними. Проте в деяких випадках доводиться знаходити ймовірність події за додаткової умови, а саме: що відбулася певна подія B. Такі ймовірності будемо називати умовними і позначати символом $P(A|B)$. На словах це означає: імовірність події A за умови, що подія B відбулася.

Приклад 1.37. Кинуто два гральних кубики. Чому дорівнює ймовірність того, що сума очок, які на них випали, дорівнює 8 (подія A), коли відомо, що ця сума є парне число (подія B). Усі можливі випадки, які можуть трапитися при киданні двох гральних кубиків, запишемо в таблицю 1.8.

(i,j)	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Таблиця 1.8

Кожна клітинка таблиці 1.8 містить запис можливої події: на першому місці в дужках ставиться число очок, що випали на першому кубику, на другому – число очок, що випали на другому кубику. Загальне число можливих випадків дорівнює 36, сприятливих події A дорівнює 5. Отже, безумовна ймовірність

$$P\left(A\right) = \frac{5}{36}.$$

Якщо подія B відбулася, то відбулася одна із 18 (а не 36) можливостей, і отже, умовна ймовірність дорівнює

$$P\left(A|B\right) = \frac{5}{18}$$

.

Приклад 1.38. Із колоди карт послідовно вийняли дві карти. Знайти:

а) безумовну ймовірність того, що друга карта буде тузом (невідомо, яка карта була вийнята першою),

б) умовну ймовірність при другому вийманні витягти туза, якщо першим витягли туза.

Позначимо через A подію, що полягає в появі туза на другому місці, а через B – подію, що полягає в появі туза на першому місці. Зрозуміло, що має місце рівність:

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

Унаслідок несумісності подій $A \cap B$ і $A \cap \overline{B}$ маємо:

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

При вийманні двох карт із колоди в 36 карт може відбутися $36 \cdot 35$ (ураховуючи порядок) випадків. Із них сприятливих події $A \cap B$ є $4 \cdot 3$ випадки і події $A \cap \overline{B}$ є $32 \cdot 4$ випадки. Отже,

$$P\left(A\right) = \frac{4 \cdot 3}{35 \cdot 36} + \frac{32 \cdot 4}{36 \cdot 35} = \frac{1}{9}.$$

Якщо при першому вийманні витягли туза, то в колоді залишилося 35 карт і серед них лише 3 тузи. Отже,

$$P\left(A|B\right) = \frac{3}{35}.$$

Загальне розв'язання задачі знаходження умовної ймовірності для класичного означення ймовірності не створює трудність. Справді, нехай із n можливих, несумісних і рівноможливих подій

$$A_1, A_2, \ldots, A_n$$

події A сприяє m подій, події B сприяє k подій, події $A \cap B$ сприяє r подій (зрозуміло, що $(r \leq k, r \leq m)$. Якщо подія B відбулася, то це означає, що настала одна з подій A_j, яка сприяє B. За цієї умови події A сприяє r і лише r подій A_j, що сприяють $A \cap B$. Отже,

$$P\left(A|B\right) = \frac{r}{k}.$$

Цю рівність можна записати інакше, якщо зауважити, що

$$P\left(B\right) = \frac{k}{n}, \quad P\left(A \cap B\right) = \frac{r}{n}.$$

Справді, легкий підрахунок показує, що

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.13}$$

Так само можна вивести, що

$$P(B|A) = \frac{P(A \cap B)}{P(A)}. \tag{1.14}$$

Зрозуміло, коли B (або A) є неможлива подія, то рівність (1.13) (відповідно і (1.14)) втрачає сенс. Кожна із рівностей (1.13), (1.14) еквівалентна так званій теоремі множення, згідно з якою

$$P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B). \tag{1.15}$$

Словами цю рівність можна виразити так:

Імовірність перетину двох подій дорівнює добутку ймовірності однієї із цих подій на умовну ймовірність другої, за умови, що перша відбулася.

Теорема множення може бути застосована й тоді, коли одна із подій A або B є неможлива подія, оскільки в цьому випадку разом з $P(A) = 0$ мають місце рівності $P(A|B) = 0$ і $P(A \cap B) = 0$.

1.5.1 Аксіоматичне означення та властивості

При аксіоматичному означенні ймовірності за основу беруть співвідношення встановлені при класичному означенні ймовірності.

Означення 1.11. Нехай $A, B \in \mathfrak{F}$ і $\mathbb{P}(B) > 0$. Умовною ймовірністю події A за умови події B називається відношення ймовірності їх перетину до ймовірності умови:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Теорема 1.19. (Теорема про властивості умовної ймовірності.)

Нехай $A, B \in \mathfrak{F}$ *і* $\mathbb{P}(B) > 0$. *Тоді*

(а) $\mathbb{P}(A|B) \geq 0$,

(б) $\mathbb{P}(\Omega|B) = \mathbb{P}(B|B) = 1$,

(в) $\mathbb{P}(\cup_n A_n|B) = \sum_n \mathbb{P}(A_n|B)$ *для попарно несумісних* $A_n \in \mathfrak{F}$.

Доведення. (а) Невід'ємність є наслідком означення та невід'ємності ймовірності.

(б) Нормованість випливає з рівності $\Omega \cap B = B$.

(в) За означенням умовної ймовірності внаслідок сигма-адитивності ймовірності з попарної несумісності подій $A_n \cap B$ отримуємо:

$$\mathbb{P}(\cup_n A_n | B) = \frac{\mathbb{P}((\cup_n A_n) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\cup_n (A_n \cap B)}{\mathbb{P}(B)} =$$

$$\sum_n \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_n \mathbb{P}(A_n | B)$$

□

Наслідок 1.2. *Умовна ймовірність $\mathbb{P}(A|B)$ як функція події A має всі основні властивості ймовірностей, що випливають з аксіом та викладені в теоремі про основні властивості ймовірностей. Дійсно, твердження (а),(б),(в) теореми повністю відповідають аксіомам теорії ймовірностей (P1),(P2),(P3).*

Теорема 1.20. (Теорема про ймовірність перерізу подій.) *Нехай $A, B \in \mathfrak{F}$ причому $\mathbb{P}(B) > 0$. Тоді*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B).$$

Зауваження 1.8. У випадку, коли $\mathbb{P}(B) > 0$, наведена формула також справедлива, оскільки за монотонністю ймовірності $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$, а права частина завжди нульова. Тому доцільно визначити $\mathbb{P}(A|B) = 0$ у випадку, коли $\mathbb{P}(B) = 0$.

Теорема 1.21. (Теорема про ймовірність перерізу n подій.) *Нехай $A_k \in \mathfrak{F}$, $k = 0, 1, \ldots, n$ такі, що $\mathbb{P}(A_k) > 0$, $k = 0, 1, \ldots, n$. Тоді*

$$\mathbb{P}(\cap_{0 \leq k \leq n} A_k) = \mathbb{P}(A_0) \prod_{1 \leq k \leq n} \mathbb{P}(A_k | A_0 \cap \cdots \cap A_{k-1}).$$

Доведення. Формула виводиться з теореми про ймовірність перерізу подій за індукцією.

□

Наслідок 1.3. *Нехай $A_1, A_2, A_3 \in \mathfrak{F}$ такі, що $\mathbb{P}(A_k) > 0$, $k = 1, 2, 3$. Тоді*

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2).$$

Такі формули використовують для аналізу складних стохастичних експериментів. Спочатку задають імовірності проміжних результатів та умовні ймовірності заключних результатів експерименту, а вже через них обчислюють сумісні ймовірності всіх подій.

1.6 Формула повної ймовірності

Означення 1.12. Послідовність подій $\{H_k, k \geq 1\}$ називається повною групою подій (гіпотез), якщо

а) $H_i \cap H_j = \emptyset$ при $i \neq j$, та

б) $\cup_k H_k = \Omega$.

Еквівалентне означення полягає в тому, що $\{H_k, k \geq 1\}$ утворюють розбиття простору Ω, тобто в результаті проведення стохастичного експерименту відбувається одна і тільки одна подія з повної групи:

$$\forall \omega \in \Omega \exists! n \geq 1 \colon \omega \in H_n.$$

Теорема 1.22. (Теорема (формула повної ймовірності.) *Нехай $\{H_k, k \geq 1\}$ – повна група подій. Для кожної події $A \in \mathfrak{F}$ виконується рівність*

$$\mathbb{P}(A) = \sum_{k \geq 1} \mathbb{P}(A|H_k)\mathbb{P}(H_k).$$

Доведення. Оскільки $A = \cup_{k \geq 1}(A \cap H_k)$, де події в правій частині попарно несумісні, то за властивістю сигма-адитивності та за теоремою про ймовірність перерізу подій

$$\mathbb{P}(A) = \sum_{k \geq 1} \mathbb{P}(A \cap H_k) = \sum_{k \geq 1} \mathbb{P}(A|H_k)\mathbb{P}(H_k).$$

□

Як ілюстрацію розглянемо два приклади.

Приклад 1.39. Маємо п’ять урн:

2 урни складу A_1 - по 2 білих і по 1 чорній кулі.

1 урна складу A_2 - 10 чорних куль.

2 урни складу A_3 - по 3 білих кулі та по 1 чорній.

Наугад вибирається урна і з неї наугад виймається куля. Чому дорівнює ймовірність того, що вийнята куля біла (подія B)?

Оскільки вийнята куля може бути лише з урни першого, другого або третього складу, то

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B).$$

За формулою повної ймовірності знаходимо, що

$$P(B) = P(A_1) P(B|A_1) + P(A_2) P(B|A_2) + P(A_3) P(B|A_3).$$

Але ясно, що

$$P(A_1) = \frac{2}{5},\ P(A_2) = \frac{1}{5},\ P(A_3) = \frac{2}{5},$$

$$P(B|A_1) = \frac{2}{3},\ P(B|A_2) = 0,\ P(B|A_3) = \frac{3}{4}.$$

Отже, за формулою повної ймовірності

$$P(B) = \frac{2}{5} \cdot \frac{2}{3} + \frac{1}{5} \cdot 0 + \frac{2}{5} \cdot \frac{3}{4} = \frac{17}{30}.$$

Приклад 1.40. У кого із студентів більше шансів обрати щасливий білет?

Екзамен складають n студентів, які послідовно обирають без повернення по одному білету з початкової загальної кількості m. Ця кількість містить $k \leq m$ щасливих білетів. Позначимо через $A_n, n \geq 1$, подію, що полягає у виборі щасливого білета n-м студентом. Тоді кожна з пар подій $(A_n, \overline{A}_n), n \geq 1$, є повною групою подій. За класичним означенням імовірностей $\mathbb{P}(A_1) = \frac{k}{m}$. Тому за формулою повної ймовірності

$$\mathbb{P}(A_2) = \mathbb{P}(A_2|A_1)\mathbb{P}(A_1) + \mathbb{P}(A_2|\overline{A}_1)\mathbb{P}(\overline{A}_1) =$$

$$= \frac{k-1}{m-1}\frac{k}{m} + \frac{k}{m-1}\frac{m-k}{m} = \frac{k}{m}.$$

За індукцією доводимо, що $\mathbb{P}(A_j) = \frac{k}{m}$ при $j \leq m$.

Вправа 1.5. У кожній з m урн знаходяться k білих та $n-k$ чорних куль. З першої урни навмання обрали кулю та переклали у другу, потім випадково обрану кулю з другої урни переклали у третю і так далі. Знайти ймовірність витягнути білу кулю з останньої урни.

1.7 Формула Байєса

Ми можемо тепер вивести важливу **формулу Байєса**. Ця формула дає можливість послідовно уточнювати ймовірності повної групи подій за результатами спостережень на підставі проміжних результатів стохастичного експерименту.

Теорема 1.23. (Формула Байєса.)
Нехай $\{H_k, k \geq 1\}$ *– повна група подій,* $A \in \mathfrak{F}$ *та* $\mathbb{P}(A) > 0$*. Тоді*

$$\mathbb{P}(H_k|A) = \frac{\mathbb{P}(A|H_k)\mathbb{P}(H_k)}{\sum_{j\geq 1} \mathbb{P}(A|H_j)\mathbb{P}(H_j)}.$$

Доведення. За означенням умовної ймовірності та за формулою повної ймовірності

$$\mathbb{P}(H_k|A) = \frac{\mathbb{P}(A \cap H_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|H_k)\mathbb{P}(H_k)}{\sum_{j\geq 1} \mathbb{P}(A|H_j)\mathbb{P}(H_j)}.$$

□

Ці формули мають назву формул Байєса. Загальна схема, за якою застосовуються ці формули до розв'язання практичних задач, така. Нехай подія B може відбуватися в різних умовах, відносно характеру яких може бути зроблено n несумісних гіпотез:

$$H_1, H_2, \ldots, H_n.$$

Нам відомі ймовірності $P(H_k)$ цих гіпотез до випробовування. Маємо також, що при гіпотезі H_k подія B має ймовірність $P(B|H_k)$. Нехай відомо, що подія B відбулася. Це повинно викликати переоцінку ймовірностей гіпотез H_k; формули Байєса розв'язують це питання.

Приклад 1.41. Маємо 5 урн такого складу:

2 урни (склад H_1) по 2 білих і по 3 чорних кулі;

2 урни (склад H_2) по 1 білій і по 4 чорних кулі;

1 урна (склад H_3) - 4 білих і 1 чорна куля.

З однієї наугад вибраної урни взято кулю. Ця куля виявилася біла (подія B). Чому дорівнює після досліду ймовірність (апостеріорна ймовірність) того, що кулю взято з урни третього складу?

Згідно з припущенням

$$P(H_1) = \frac{2}{5},\ P(H_2) = \frac{2}{5},\ P(H_3) = \frac{1}{5},$$

$$P\left(B|H_1\right)=\frac{2}{5},\ P\left(B|H_2\right)=\frac{1}{5},\ P\left(B|H_3\right)=\frac{4}{5}.$$

Згідно з формулою Байєса маємо

$$P\left(H_3|B\right)=\frac{P\left(H_3\right)P\left(B|H_3\right)}{P\left(H_1\right)P\left(B|H_1\right)+P\left(H_2\right)P\left(B|H_2\right)+P\left(H_3\right)P\left(B|H_3\right)}=$$

$$=\frac{\frac{1}{5}\cdot\frac{4}{5}}{\frac{2}{5}\cdot\frac{2}{5}+\frac{2}{5}\cdot\frac{1}{5}+\frac{1}{5}\cdot\frac{4}{5}}=\frac{4}{10}=\frac{2}{5}.$$

Так само знаходимо

$$P\left(H_1|B\right)=\frac{2}{5},\ P\left(H_2|B\right)=\frac{1}{5}.$$

Приклад 1.42. Оцінка невідомого складу урни за кольором навмання обраної кулі.

Урна містить n куль. Кількість білих куль в урні невідома. Всі апріорні припущення щодо вмісту урни рівноможливі. Позначимо через H_k, $k=0,1,\dots,n$, гіпотезу про те що в урні k, $k=0,1,\dots,n$, білих куль, інші – чорні. Оскільки всі припущення щодо вмісту урни рівноможливі то ймовірності гіпотез рівні. Отже

$$\mathbb{P}(H_k)=\frac{1}{n+1}.$$

Умовні ймовірності вибрати білу кулю (подія A) легко обчислюються

$$\mathbb{P}(A|H_k)=\frac{k}{n}.$$

З урни навмання обрано кулю, яка виявилася білою. Як треба змінити ймовірності гіпотез щодо складу урни? Нехай A – подія, що полягає у виборі білої кулі. За умовою рівноможливості гіпотез

$$\mathbb{P}(H_k|A)=\frac{\mathbb{P}(A|H_k)\mathbb{P}(H_k)}{\sum_{j=0}^{n}\mathbb{P}(A|H_j)\mathbb{P}(H_j)}=$$

$$\frac{k}{n}\frac{1}{n+1}\left(\sum_{j=0}^{n}\frac{j}{n}\frac{1}{n+1}\right)^{-1}=\frac{2k}{n(n+1)}.$$

Вправа 1.6. (Льюіс Керрол) В урні знаходиться одна куля невідомого кольору. В урну поклали білу кулю. Після цього наудачу обрана з урни куля виявилася білою. Яка ймовірність того, що спочатку в урні була біла куля?

1.8 Незалежні випадкові події

У теорії ймовірностей використовується специфічне поняття незалежності, яке можна назвати статистичною незалежністю. Дві події вважають незалежними, якщо наявна інформація про одну з них не змінює шансів для іншої.

Як вказував А.М. Колмогоров, насамперед поняття незалежності відрізняє предмет теорії ймовірностей від теорії міри.

1.8.1 Незалежні події та їх властивості

Означення 1.13. Нехай $A, B \in \mathfrak{F}$ причому $\mathbb{P}(B) > 0$. Подія A не залежить від події B, якщо

$$\mathbb{P}(A|B) = \mathbb{P}(A).$$

Якщо подія A незалежна від B, то за рівністю (1.15) має місце рівність

$$\mathbb{P}(A) \cdot \mathbb{P}(B|A) = \mathbb{P}(B) \cdot \mathbb{P}(A|B).$$

Звідси знаходимо, що

$$\mathbb{P}\left(B|A\right) = \mathbb{P}\left(B\right), \tag{1.16}$$

тобто, що подія B також незалежна від A. Отже, властивість незалежності подій взаємна.

З означення умовної ймовірності отримуємо еквівалентне, але більш загальне та симетричне означення попарної незалежності, яке відоме як **теорема множення**.

Означення 1.14. Дві події $A, B \in \mathfrak{F}$ називаються незалежними, якщо

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

Зауваження 1.9. У практичних задачах для визначення незалежності даних подій рідко звертаються до перевірки виконання для них вказаних рівностей. Зазвичай для цього користуються міркуваннями, заснованими на досвіді.

Приклад 1.43. Підкидання двох монет. Простір всіх елементарних подій $\Omega = \{AA, AP, PA, PP\}$ складається із 4 елементарних подій, $N(\Omega) = 4$. Події B {при підкиданні двох монет першим випаде A} відповідає підмножина $B = \{AA, AP\}$ простору Ω із 2 елементарних подій, $N(B) = 2$. Події C {при підкиданні двох монет другий раз випаде A} відповідає підмножина $C = \{AA, PA\}$ простору Ω із 2 елементарних подій, $N(C) = 2$. Події $B \cap C$ {при першому і при другому підкиданні двох монет випаде A} відповідає підмножина $B \cap C = \{AA\}$ простору Ω із 1 елементарної події, $N(B \cap C) = 1$. Тому

$$\mathbb{P}(B) = \frac{2}{4} = \frac{1}{2}, \quad \mathbb{P}(C) = \frac{2}{4} = \frac{1}{2}, \quad \mathbb{P}(B \cap C) = \frac{1}{4}.$$

Маємо

$$\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C).$$

Отже події B та C незалежні.

Приклад 1.44. Розглянемо тепер результати підкидання двох кубиків. Випадання кожної із 36 можливих комбінацій числа очок на першому і на другому кубиках вважається рівноможливим. Простір всіх елементарних подій Ω складається із 36 елементарних подій, $N(\Omega) = 36$. Події B {при підкиданні двох кубиків перший раз випаде шістка} відповідає підмножина $B = \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$ простору Ω із 6 елементарних подій, $N(B) = 6$. Події C {при підкиданні двох кубиків другий раз випаде шістка} відповідає підмножина $C = \{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6)\}$ простору Ω із 6 елементарних подій, $N(C) = 6$. Події $B \cap C$ {при першому і при другому підкиданні двох кубиків випаде шістка} відповідає підмножина $B \cap C = \{(6,6)\}$ простору Ω із 1 елементарної події, $N(B \cap C) = 1$. Тому

$$\mathbb{P}(B) = \frac{6}{36} = \frac{1}{6}, \quad \mathbb{P}(C) = \frac{6}{36} = \frac{1}{6}, \quad \mathbb{P}(B \cap C) = \frac{1}{36}.$$

Маємо

$$\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C).$$

Отже події B та C незалежні.

Теорема 1.24. (Теорема про перетворення незалежних подій.)

(а) Довільна подія A не залежить від Ω та $\varnothing$.

(б) Якщо події A і B B незалежні, то $\overline{A}$ і B, A і $\overline{B}$, $\overline{A}$ і $\overline{B}$ також незалежні.

(в) Якщо подія A не залежить від B_1 та від B_2, причому $B_1 \subset B_2$, то A і $B_2 \setminus B_1$ також незалежні.

(г) Якщо подія A не залежить від подій $B_n, n \geq 1$, що є попарно несумісними: $B_i \cap B_j = \varnothing, \forall i \neq j$, то події A і $\cup_{n\geq 1} B_n$ також незалежні

Доведення. (а)

$$\mathbb{P}(A \cap \Omega) = \mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(\Omega).$$

(б)

$$\mathbb{P}(\overline{A} \cap B) = \mathbb{P}(B \setminus (A \cap B)) = \mathbb{P}(B) - \mathbb{P}(A \cap B) =$$

$$= \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(\overline{A})\mathbb{P}(B).$$

(в)

$$\mathbb{P}(A \cap (B_2 \setminus B_1)) = \mathbb{P}((A \cap B_2) \setminus (A \cap B_1)) =$$

$$= \mathbb{P}(A \cap B_2) - \mathbb{P}(A \cap B_1) = \mathbb{P}(A)\mathbb{P}(B_2) - \mathbb{P}(A)\mathbb{P}(B_1) =$$

$$= \mathbb{P}(A)(\mathbb{P}(B_2) - \mathbb{P}(B_1)) = \mathbb{P}(A)(\mathbb{P}(B_2 \setminus B_1).$$

(г)

$$\mathbb{P}(A \cap (\cup_{n\geq 1} B_n)) = \mathbb{P}(\cup_{n\geq 1}(A \cap B_n)) = \sum_{n\geq 1} \mathbb{P}(A \cap B_n) =$$

$$\sum_{n\geq 1} \mathbb{P}(A)\mathbb{P}(B_n) = \mathbb{P}(A) \sum_{n\geq 1} \mathbb{P}(B_n) = \mathbb{P}(A)\mathbb{P}(\cup_{n\geq 1} B_n).$$

□

1.8.2 Незалежність подій в сукупності, приклад Бернштейна

Поняття незалежності можна поширити на послідовності подій.

Означення 1.15. Події з послідовності $\{A_k, k = 1, \ldots, n\}$ називаються попарно незалежними, якщо для всіх $i \neq j$ події A_i та A_j незалежні.

Означення 1.16. Події з послідовності $\{A_k, k = 1, \ldots, n\}$ називаються незалежними в сукупності, якщо для довільних натуральних $m, 1 \leq m \leq n$ і довільних натуральних $1 \leq k_1 \leq k_1 \leq \cdots \leq n$

$$\mathbb{P}(A_{k_1} \cap A_{k_2} \cap \cdots \cap A_{k_m}) = \mathbb{P}(A_{k_1})\mathbb{P}(A_{k_2}) \cdots \mathbb{P}(A_{k_m}).$$

Очевидно, що з незалежності в сукупності випливає попарна незалежність подій послідовності $\{A_k, k = 1, \ldots, n\}$.

Наступний приклад С.Н.Бернштейна доводить що з попарної незалежності не випливає незалежність у сукупності.

Приклад 1.45. Нехай маємо правильний тетраедр із 4 гранями. Уявімо собі, що грані тетраедра пофарбовані: 1-ша в червоний колір (A), 2-га – у зелений (B), 3-тя – у синій (C) і 4-та в усі три кольори ($A \cap B \cap C$). Тетраедр навмання підкинули і він впав на одну з граней. Легко бачити, що ймовірність того, що грань, на яку впаде тетраедр при киданні, матиме червоний колір, дорівнює $1/2$; граней – чотири і дві з них забарвлені у червоний колір. Отже,

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

Так само можна підрахувати, що

$$\mathbb{P}(B) = \mathbb{P}(C) = \mathbb{P}(A|B) = \mathbb{P}(B|C) = \mathbb{P}(C|A) = \mathbb{P}(B|A) = \\ = \mathbb{P}(C|B) = \mathbb{P}(A|C) = \tfrac{1}{2},$$

події A, B, C, таким чином, попарно незалежні. Проте, якщо нам відомо, що відбулися події B і C, то напевно відбулася і подія A, тобто

$$\mathbb{P}(A|B \cap C) = 1.$$

Отже, події A, B, C, в сукупності залежні.

2 Випадкові величини та функції розподілу

2.1 Загальне означення випадкової величини та вектора

Для формулювання загального означення слід структурувати клас множин у просторі значень випадкової величини.

2.1.1 Борелєва сигма-алгебра, борелєві множини

Розглянемо на числовій осі $\mathbb{R}$ клас напівінтервалів

$$\mathfrak{J}(\mathbb{R}) = \{[a, b), -\infty \le a < b \le \infty\}.$$

Скінченні об'єднання попарно несумісних напівінтервалів є алгеброю підмножин $\mathbb{R}$:

$$\mathfrak{U}(\mathbb{R}) = \{A = \cup_{k=1}^{n}[a_k, b_k), -\infty \le a_1 < b_1 < a_2 < b_2 < \cdots < b_n \le \infty\}.$$

Нагадаємо, що переріз будь-якого класу σ–алгебр завжди є σ-алгеброю, яка називається найменшою σ-алгеброю з даного класу.

Означення 2.1. Борелєвою сигма-алгеброю $\mathfrak{B}(\mathbb{R})$ називається найменша σ-алгебра, яка містить алгебру $\mathfrak{U}(\mathbb{R})$, (або клас $\mathfrak{J}(\mathbb{R})$, що еквівалентно).

Інакше кажучи, σ-алгебра $\mathfrak{B}(\mathbb{R})$ породжена алгеброю $\mathfrak{U}(\mathbb{R})$: $\mathfrak{B}(\mathbb{R}) = \sigma[\mathfrak{U}(\mathbb{R})]$.

Будь-який елемент $B \in \mathfrak{B}(\mathbb{R})$ називається борелєвою множиною, або ж борелівською множиною.

Аналогічно, система n-вимірних кутів

$$\mathfrak{J}(\mathbb{R}^n) = \left\{A = (-\infty, a) = \prod_{k=1}^{n}(-\infty, a_k), a_k \in \mathbb{R} \cup \infty\right\}$$

у n-вимірному просторі породжує алгебру $\mathfrak{U}(\mathbb{R}^n)$ скінченних об'єднань напіввідкритих паралелепіпедів, що не перетинаються.

Означення 2.2. Борелєвою σ-алгеброю $\mathfrak{B}(\mathbb{R}^n)$ у $\mathbb{R}^n$ називається найменша сигма-алгебра, яка містить алгебру $\mathfrak{U}(\mathbb{R}^n)$.

Означення 2.3. Функція $g : \mathbb{R}^k \to \mathbb{R}^n$ називається борелєвою, якщо
$$\{x : g(x) \in B_n\} \in \mathfrak{B}(\mathbb{R}^k), \forall B_n \in \mathfrak{B}(\mathbb{R}^n).$$

Зокрема, всі неперервні функції є борелєвими. Дійсно, прообраз відкритого інтервалу для неперервної функції є відкритою множиною – отже, борелєвою. Оскільки клас множин, прообразами яких є борелєві множини, є σ-алгеброю та містить всі відкриті напівінтервали $(-\infty, a)$, то цей клас збігається з борелєвою σ-алгеброю.

2.1.2 Загальне означення випадкової величини

Означення 2.4. Випадковою величиною називається така числова функція $\xi : \Omega \to \mathbb{R}$, що для довільної борелєвої множини $B \in \mathbb{R}$, її прообраз: $\{\omega : \xi(\omega) \in B\}$ є випадковою подією, тобто:

$$\{\omega : \xi(\omega) \in B\} \in \mathfrak{F},\ \forall B \in \mathfrak{B}(\mathbb{R}).$$

Функції ξ, які задовольняють наведену умову, називаються вимірними відносно сигма-лгебри $\mathfrak{F}$. Отже, для випадкових величин (вимірних функцій) можна говорити про ймовірність події

$$\{\xi \in B\} = \{\omega : \xi(\omega) \in B\},\ \forall B \in \mathfrak{B}(\mathbb{R}).$$

Теорема 2.1. (Критерій вимірності скалярної функції.) *Нехай $\xi : \Omega \to \mathbb{R}$ - довільна функція. Вона є випадковою величиною тоді й тільки тоді, коли прообраз кожного напівінтервалу $(-\infty, x)$ є випадковою подією:*

$$\{\xi < x\} = \{\omega : \xi(\omega) < x\} \in \mathfrak{F},\ \forall x \in \mathbb{R}.$$

Цю властивість вимірних функцій можна використати щоб дати таке означення випадкової величини.

Означення 2.5. Випадковою величиною називається така числова функція $\xi : \Omega \to \mathbb{R}$, що для кожного напівінтервалу $(-\infty, x), x \in \mathbb{R}$, його прообраз $\{\xi < x\} = \{\omega : \xi(\omega) < x\}$ є випадковою подією, тобто:

$$\{\xi < x\} = \{\omega : \xi(\omega) < x\} \in \mathfrak{F},\ \forall x \in \mathbb{R}.$$

Теорема 2.2. (Вимірність суперпозиції.) *Нехай $\xi : \Omega \to \mathbb{R}$ – випадкова величина, а $g : \mathbb{R} \to \mathbb{R}$ – борелєва функція. Тоді суперпозиція $\eta = g(\xi)$ є випадковою величиною.*

Доведення. Обчислимо

$$\{\eta < x\} = \{g(\xi) < x\} = \{\xi \in g^{(-1)}((-\infty, x))\} \in \mathfrak{F},$$

оскільки $g^{(-1)}((-\infty, x))$ – борелєва множина за означенням функції g. □

Означення 2.6. Випадковим вектором у просторі $\mathbb{R}^n$ називається така функція $\xi : \Omega \to \mathbb{R}^n$, що для довільної борелєвої множини $B \in \mathfrak{B}(\mathbb{R}^n)$ її прообраз $\{\omega : \xi(\omega) \in B\}$ є випадковою подією:

$$\{\omega : \xi(\omega) \in B\} \in \mathfrak{F}, \forall B \in \mathfrak{B}(\mathbb{R}^n).$$

Як і для випадкових величин, можна дати таке означення випадкового вектора.

Означення 2.7. Випадковим вектором називається така функція $\xi : \Omega \to \mathbb{R}^n$, що для кожного кута $(-\infty, x) = \prod_{k=1}^{n}(-\infty, x_k)$ його прообраз є вимірною множиною

$$\{\xi < x\} = \{\omega : \xi_1(\omega) < x_1, \ldots, \xi_n(\omega) < x_n\} \in \mathfrak{F}, x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Теорема 2.3. (Вимірність функції від випадкового вектора.) *Нехай $\xi : \Omega \to \mathbb{R}^n$ – випадковий вектор у $\mathbb{R}^n$, а $g : \mathbb{R}^n \to \mathbb{R}^m$ – борелева функція. Тоді суперпозиція $g(\xi)$ є випадковим вектором у $\mathbb{R}^m$.*

Теорема 2.4. (Перетворення випадкових величин.) *Якщо ξ, η – випадкові величини, то*

$$|\xi|,\ \xi + \eta,\ \xi - \eta,\ \xi \cdot \eta,\ \xi/\eta,\ \max\{\xi, \eta\}$$

також є випадковими величинами.

Доведення. Доведення випливає з неперервності (отже, борелєвості) функцій

$$|x|,\ x + y,\ x - y,\ x \cdot y,\ x/y,\ \max\{x, y\}$$

від випадкового вектора (ξ, η). □

2.2 Функція розподілу. Основні властивості

Означення 2.8. Функцією розподілу випадкової величини $\xi : \Omega \to \mathbb{R}$ називається функція $F_\xi(x)$ дійсного аргументу $x \in \mathbb{R}$, яка задається рівністю

$$F_\xi(x) = \mathbb{P}\{\xi < x\} = \mathbb{P}\{\omega : \xi(\omega) < x\},\ x \in \mathbb{R}.$$

Теорема 2.5. (Основні властивості функції розподілу.) *Нехай випадкова величина $\xi : \Omega \to \mathbb{R}$ має функцію розподілу $F(x) = F_\xi(x) = \mathbb{P}\{\xi < x\}$. Тоді ця функція*

(1) неспадна: $F(x) \le F(y)$ для всіх $x \le y$,

(2) нормована: $F(-\infty) = 0$, $F(+\infty) = 1$,

(3) неперервна зліва: $F(x-0) = F(x)$ для всіх $x \in \mathbb{R}$.

Зауваження 2.1. Тут та надалі символом $F(x-0)$ позначається монотонна границя

$$F(x-0) = \lim_{y\uparrow x, y<x} F(y).$$

Аналогічно визначаються границі

$$F(-\infty) = \lim_{y\downarrow -\infty} F(y), \quad F(+\infty) = \lim_{y\uparrow +\infty} F(y).$$

Існування та єдиність границь у (2) та (3) випливає з монотонності (1) функції розподілу.

Означення 2.9. Альтернативне означення функції розподілу має вигляд

$$F_\xi(x) = \mathbb{P}\{\xi \le x\} = \mathbb{P}\{\omega : \xi(\omega) \le x\},$$

що відрізняється властивістю неперервності справа.

Доведення. (1) Нерівність є наслідком монотонності ймовірності, тому що при $x < y$ має місце включення $\{\xi < x\} \subset \{\xi < y\}$.

(2) Границя монотонної функції не залежить від вибору апроксимуючої послідовності аргументів. Тому y можна вважати цілозначними. Оскільки послідовність подій $\{\xi < n\}$ монотонно не зростає при $n \downarrow -\infty$, а $\cap_{n<0}\{\xi < n\} = \emptyset$, то $\{\xi < n\}\downarrow \emptyset$, Отже, за неперервністю ймовірності

$$F(-\infty) = \lim_{n\downarrow -\infty} F(n) = \lim_{n\downarrow -\infty} \mathbb{P}\{\xi < n\} = \mathbb{P}\{\emptyset\} = 0.$$

Аналогічно, з того що $\{\xi < n\}\uparrow\Omega$ при $n\uparrow\infty$ дістанемо

$$F(\infty) = \lim_{n\uparrow\infty} F(n) = \lim_{n\uparrow\infty} \mathbb{P}\{\xi < n\} = \mathbb{P}\{\Omega\} = 1.$$

(3) Клас відкритих напівінтервалів $(-\infty, x), x \in \mathbb{R}$, є неперервним зліва для монотонної збіжності: $(-\infty, y) \uparrow (-\infty, x)$ при $y \uparrow x$, оскільки $(-\infty, x) = \cup_{n\geq 1}(-\infty, x - 1/n)$. Тому має місце монотонна збіжність при $y \uparrow x$

$$\{\xi \in (-\infty, y)\} = \{\xi < y\} \uparrow \{\xi < x\} = \{\xi \in (-\infty, x)\}$$

Отже, за неперервністю ймовірностей

$$F(x-0) = \lim_{n\to\infty} F(x-1/n) = \lim_{n\to\infty} \mathbb{P}\{\xi < x-1/n\} = \mathbb{P}\{\xi < x\} = F(x).$$

□

Теорема 2.6. (Властивості функції розподілу.) *Нехай випадкова величина $\xi : \Omega \to \mathbb{R}$ має функцію розподілу $F(x) = F_\xi(x)$. Тоді*

(а) $\mathbb{P}\{a \leq \xi < b\} = F_\xi(b) - F_\xi(a)$,
(б) $\mathbb{P}\{\xi = x\} = F_\xi(x + 0) - F_\xi(x)$,
(в) $\mathbb{P}\{\xi \geq x\} = 1 - F_\xi(x)$,
(г) $\mathbb{P}\{\xi \leq x\} = F_\xi(x + 0)$,
(д) $\mathbb{P}\{\xi > x\} = 1 - F_\xi(x + 0)$.

Доведення. (а) З формули для ймовірності вкладеної різниці подій отримуємо

$$\mathbb{P}\{a \leq \xi < b\} = \mathbb{P}\{(\xi < b) \setminus (\xi < a)\} = F_\xi(b) - F_\xi(a),$$

(б) Зі співвідношення

$$\{\xi = x\} = \cap_{n>1}\{x \leq \xi < x + 1/n\},$$

неперервності ймовірності та твердження (а) дістанемо

$$\mathbb{P}\{\xi = x\} = \lim_{n\to\infty} \mathbb{P}\{x \leq \xi < x + 1/n\} =$$

$$= \lim_{n\to\infty} F_\xi(x + 1/n) - F_\xi(x) = F_\xi(x + 0) - F_\xi(x),$$

(в) Є наслідком означення функції розподілу та формули про ймовірність доповнення.

(г),(д) Виводяться з означення функції розподілу та (б).

□

Зауваження 2.2. Будь-яка функція $F : \mathbb{R} \to \mathbb{R}$, яка задовольняє умови (1), (2) і (3), є функцією розподілу деякої випадкової величини.

Означення 2.10. Нехай $F(x)$ – довільна функція розподілу. Назвемо адитивною мірою Лебега – Стілтьєса, породженою F (або, коротше, адитивною мірою F), таку функцію F на алгебрі $\mathfrak{U}(\mathbb{R})$ підмножин $\mathbb{R}$ вигляду

$$A = \cup_{k=1}^{n}[a_k, b_k), \quad -\infty \leq a_1 < b_1 < a_2 < \cdots < b_n \leq \infty,$$

значення якої задається рівністю

$$F(A) = \sum_{k=1}^{n}(F(b_k) - F(a_k))$$

.

Теорема 2.7. (Про адитивну міру Лебега – Стілтьєса.) *Адитивна міра F є невід'ємною, нормованою та скінченно-адитивною функцією на $\mathfrak{U}(\mathbb{R})$. Зокрема, F напівадитивна та монотонна.*

Доведення. Невід'ємність є очевидним наслідком монотонності (невід'ємності приростів) функції розподілу F, нормованість виводиться з нормованості F. Для доведення адитивності зауважимо, що з несумісності множин $A, B \in \mathfrak{U}(\mathbb{R})$ випливає попарна несумісність інтервалів, які складають ці множини. Тому сума приростів F на інтервалах з об'єднання $A \cup B$ перегрупуванням зводиться до суми відповідних сум для A та B, що і доводить адитивність. Згідно з теоремою про основні властивості ймовірностей напівадитивність та монотонність є наслідками невід'ємності та адитивності. □

Теорема 2.8. (Неперервність у нулі міри Лебега – Стілтьєса.) *Адитивна міра Лебега – Стілтьєса F є неперервною в нулі на $\mathfrak{U}(\mathbb{R})$ і, як наслідок, сигма-адитивною на $\mathfrak{U}(\mathbb{R})$.*

Еквівалентність неперервності в нулі та сигма-адитивності для адитивних мір доведена у теоремі про основні властивості ймовірностей.

Теорема 2.9. (Про існування міри Лебега – Стілтьєса.) *Нехай F – довільна функція розподілу. На борелєвій сигма-алгебрі $\mathfrak{B}(\mathbb{R})$ існує єдина невід'ємна, нормована та сигма-адитивна міра F така, що*

$$F((-\infty, x)) = F(x), \quad \forall x \in \mathbb{R}.$$

Доведення. За теоремою про адитивну міру Лебега – Стілтьєса побудуємо адитивну ймовірність F на алгебрі, як це зроблено вище. На підставі її σ-адитивності на алгебрі $\mathfrak{U}(\mathbb{R})$ застосуємо теорему Каратеодорі про продовження міри з алгебри на породжену σ-алгебру $\mathfrak{B}(\mathbb{R})$. □

Означення 2.11. Імовірнісна міра F на $\mathfrak{B}(\mathbb{R})$, що побудована за останньою теоремою, називається **мірою Лебега – Стілтьєса**, яка породжена функцією розподілу F.

Теорема 2.10. (Обчислення ймовірностей, пов'язаних із випадковою величиною.) *Нехай випадкова величина ξ має функцію розподілу F_ξ. Тоді для довільної борелевої множини $B \in \mathfrak{B}(\mathbb{R})$, виконується рівність*

$$\mathbb{P}(\xi \in B) = F_\xi(B),$$

де F_ξ - міра Лебега - Стілтьєса, що породжена функцією розподілу F_ξ.

Доведення. За означенням сформульована рівність має місце для напівінтервала $B = (-\infty, x)$. З іншого боку, вирази в обох частинах як функції множини B є сигма-адитивними мірами. Для лівої частини це випливає з властивостей прообразу для відображення ξ та σ-адитивності ймовірності $\mathbb{P}$, а для правої частини виконується за означенням. Тому за теоремою Каратеодорі про продовження міри вказані міри збігаються на всій породженій σ-алгебрі $\mathfrak{B}(\mathbb{R})$, □

Враховуючи останню теорему, розподілом випадкової величини ξ будемо називати відповідну міру Лебега – Стілтьєса F_ξ на σ-алгебрі $\mathfrak{B}(\mathbb{R})$,

Теорема 2.11. (Обчислення розподілу функції від випадкової величини.) *Якщо випадкова величина ξ має функцію розподілу $F_\xi(x)$, а функція $g : \mathbb{R} \to \mathbb{R}$ – борелева, то для всіх $B \in \mathfrak{B}(\mathbb{R})$*

$$\mathbb{P}\{g(\xi) \in B\} = \mathbb{P}\{\xi \in g^{(-1)}(B)\} = F_\xi(g^{(-1)}(B)),$$

де $g^{(-1)}(B)$ - прообраз відображення g.

Доведення. Доведення є очевидним наслідком теореми про обчислення ймовірностей, пов'язаних із випадковою величиною, оскільки за означенням прообразу відображення

$$\{g(\xi) \in B\} = \{\xi \in g^{(-1)}(B)\}.$$

$\square$

Наслідок 2.1. (Розподіл лінійного перетворення випадкової величини.) *Для лінійної функції $g(x) = a + bx$, $b > 0$, функція розподілу лінійного перетворення $\zeta = a + b\xi$ має вигляд*

$$F_\zeta(x) = \mathbb{P}(\zeta < x) = \mathbb{P}(\{\xi < (x-a)/b\} = F_\xi((x-a)/b).$$

2.3 Дискретні функції розподілу

Означення 2.12. Дискретною випадковою величиною називається відображення $\xi : \Omega \to \mathbb{R}$, яке має скінченну або зліченну множину значень:

$$\xi(\Omega) = \{\xi(\omega), \omega \in \Omega\} = \{x_1, x_2, \dots, x_n, \dots\},$$

причому для кожного можливого значення x_n його прообраз: $\{\omega : \xi(\omega) = x_n\}$ є випадковою подією, тобто

$$\{\omega : \xi(\omega) = x_n\} \in \mathfrak{F}, \quad \forall x_n.$$

Остання умова (вимірність) є суттєвою, тому що при її порушенні ймовірності деяких висловлювань щодо значень ξ не були б визначеними.

Зауваження 2.3. Множина значень $\{x_n\}$ може бути числовою, чи множиною векторів, або навіть підмножиною довільного лінійного простору. У подальшому будемо вважати, що всі значення $\{x_n\}$ є різними.

2.3.1 Розподіл дискретної величини

Означення 2.13. Розподілом дискретної випадкової величини $\xi(\omega)$ зі значеннями $\{x_n\}$ називається послідовність імовірностей

$$p_n = p_\xi(x_n) = \mathbb{P}(\xi = x_n), \quad n = 1, 2, \dots$$

Ця послідовність є невід'ємною, $p_n \geq 0$, і в сумі становить 1, $\sum_{n=1}^{\infty} p_n = 1$, тобто є дискретним розподілом імовірностей.

Зауваження 2.4. Оскільки аргументом будь-якої випадкової величини $\xi(\omega)$ завжди є елементарна подія $\omega \in \Omega$, то при запису випадкової величини цей аргумент часто не вказують і просто пишуть ξ. Крім того, при запису ймовірностей подій, пов'язаних із випадковою величиною, для скорочення часто опускають вказівку на елементарну подію:

$$\{\xi = x\} = \{\omega : \xi(\omega) = x\}.$$

Зокрема, за домовленістю

$$\mathbb{P}(\xi = x) = \mathbb{P}(\{\omega : \xi(\omega) = x\}).$$

Розподіл дискретних випадкових величин зображують у такій таблиці 2.1

Розподіл дискретної величини						
ξ	x_1	x_2	x_3	$\dots$	x_n	$\dots$
$\mathbb{P}$	p_1	p_2	p_3	$\dots$	p_n	$\dots$

Таблиця 2.1

Теорема 2.12. (Розподіл функції від дискретної величини.) *Нехай* ξ*– дискретна випадкова величина з розподілом* $\{p_n = p_\xi(x_n)\}$ *та значеннями* $\{x_n\}$*. Тоді для довільної функції* g *справедлива рівність*

$$\mathbb{P}(g(\xi) \in B) = \sum_{n: g(x_n) \in B} p_\xi(x_n).$$

Зауважимо, що значення $g(x_n)$ можуть бути однаковими.

Доведення. Доведення є наслідком зображення

$$\{g(\xi) \in B\} = \cup_{n: g(x_n) \in B} \{\xi = x_n\}.$$

де події справа попарно несумісні, оскільки всі значення x_n – різні. □

Дискретна функція розподілу (функція розподілу дискретної випадкової величини) є сталою в кожному околі своєї точки неперервності та збігається з сумою своїх стрибків на інтервалі $(-\infty, x)$. Функція розподілу дискретної величини чисто розривна та кусково-стала. Дійсно, якщо величина ξ набуває значень $\{x_n, n \geq 1\}$ із імовірностями $\{p_n, n \geq 1\}$, то

$$F_\xi(x) = \mathbb{P}\left(\cup_{n:x_n<x}\{\xi = x_n\}\right) = \sum_{n:x_n<x} p_n.$$

Наявність у функції розподілу випадкової величини ξ стрибка p_n у точці x_n означає, що величина ξ може набути значення x_n з додатною імовірністю c_n.

Доведемо, що функція розподілу може мати не більше, ніж зліченну множину стрибків. Справді, функція розподілу може мати стрибків розміру від $1/2$ до 1 не більше одного; стрибків розміру від однієї чверті до $1/2$ – не більше трьох. Взагалі стрибків розміру від $1/2^n$ до $1/2^{n-1}$ – не більше ніж $2^n - 1$. Цілком очевидно, що можна перенумерувати всі стрибки, розмістивши їх за величиною, починаючи від більших значень і повторюючи рівні значення стільки разів, скільки стрибків цієї величини має функція $F(x)$.

2.3.2 Біноміальний розподіл

Позначимо через μ число появ події A в послідовності n незалежних випробовувань, у кожному з яких імовірність її появи стала і дорівнює p. Залежно від випадку μ може набувати всіх цілих значень від 0 до n включно. Елементарні події тут складаються з можливих появ і непояв події A при n випробовуваннях. Наприклад, однією з елементарних подій буде поява події A в кожному з n випробовувань. Усього, як неважко бачити, буде 2^n елементарних подій. Функцію від елементарної події ω визначаємо так: вона дорівнює кількості появ події A в елементарній події ω. Якщо позначити через A_k множину тих елементарних подій, для яких $\mu = k$, то маємо

$$\mathbb{P}\left\{\mu = k\right\} = \mathbb{P}\left\{A_k\right\} = P_n(k) = C_n^k p^k q^{n-k}.$$

Вимірність функцій $\mu = \mu(\omega)$ очевидна. Звідси згідно з означенням робимо висновок, що μ є випадкова величина. Для неї функція розподілу визначається так:

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \sum\limits_{k=0}^{[x]} P_n(k), & 0 < x \leq n, \\ 1, & x > n. \end{cases}$$

Функція розподілу являє собою східчасту лінію із стрибками в точках $x = 0, 1, 2, \ldots, n$. Стрибок у точці k дорівнює $P_n(k)$.

Даний приклад показує, що так звана схема Бернуллі може бути включена до загальної теорії випадкових величин.

Послідовність імовірностей

$$P_n(k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \ldots, n$$

називається **біноміальним розподілом** імовірностей (позначення $\mu \simeq Bi(n, p)$).

Теорема 2.13. (Властивості біноміальних імовірностей.) *Біноміальні ймовірності $P_n(k)$ не спадають при $k \leq (n+1)p$, та не зростають при $k > (n+1)p$, тому найбільш імовірне значення біноміального розподілу дорівнює $k = [(n+1)p]$.*

Доведення. Обчислимо послідовні відношення

$$\frac{P_n(k)}{P_n(k-1)} = \frac{n-k+1}{k}\frac{p}{q} = 1 + \frac{(n+1)p - k}{kq}.$$

Права частина менша за 1 тоді й тільки тоді, коли $k > (n+1)p$, звідки випливає необхідний висновок. $\square$

2.3.3 Поліноміальний розподіл

Означення 2.14. Поліноміальною схемою випробувань називається послідовність з n незалежних випробувань, результат кожного з яких належить множині $\{x_1, x_2, \ldots, x_m\}$ та має дискретний розподіл $p = \{p_i = P(\xi(k) = x_i), i = 1, 2, \ldots, m\}$.

Результатом таких випробувань є вектор

$$\nu(n) = \{\nu_1, \nu_2, \ldots, \nu_m\},$$

де ν_i - кількість випробувань із результатом x_i у n випробуваннях. З незалежності випробувань випливає формула для поліноміального розподілу

$$P_n(n_1, n_2, \ldots, n_m) = P\{\nu_1 = n_1, \nu_2 = n_2, \ldots, \nu_m = n_m\} =$$

$$= \frac{n!}{n_1! n_2! \cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}.$$

2.3.4 Геометричний розподіл

Дискретна випадкова величина ξ має геометричний розподіл із параметром p, позначення $\xi \simeq Geo(p)$, якщо вона збігається з кількістю випробувань до першого успіху в нескінченній послідовності випробувань Бернуллі з імовірністю успіху p в одному випробуванні.

Геометричний розподіл імовірностей має вигляд

$$\mathbb{P}(\xi = k) = (1-p)^{k-1}p,\ k = 1, 2, \dots$$

Іноді розглядають також величину, що дорівнює кількості неуспіхів до першого успіху, з розподілом

$$\mathbb{P}(\xi = k) = (1-p)^{k}p,\ k = 0, 1, 2, \dots$$

Геометричний розподіл ймовірностей має властивість, яка називається **відсутність післядії.** Ця властивість записується так

$$\mathbb{P}(\xi = n+m|\xi \geq n) = \mathbb{P}(\xi = m),\ m = 1, 2, \dots$$

Дійсно,

$$\mathbb{P}(\xi = n+m|\xi \geq n) = \frac{\mathbb{P}(\xi = n+m, \xi \geq n)}{\mathbb{P}(\xi \geq n)} =$$

$$= \frac{\mathbb{P}(\xi = n+m)}{\mathbb{P}(\xi \geq n)} = \frac{q^{n+m}}{\sum_{k=n}^{\infty} q^k p} = q^m p = \mathbb{P}(\xi = m).$$

Варто відмітити, що серед усіх дискретних розподілів, які приймають цілі значення $0, 1, \dots$ лише геометричний розподіл має таку властивість. Дійсно, нехай випадкова величина ξ приймає значення $0, 1, \dots, n, \dots$ з ймовірностями $p_n, n = 0, 1, \dots$. Нехай цей розподіл має властивість відсутності післядії, тобто

$$\frac{p_{n+m}}{\sum_{k=n}^{\infty} p_k} = p_m.$$

Покладемо у цій рівності $n = 1$. Оскільки $\sum_{k=0}^{\infty} p_k = 1$, то матимемо

$$p_{m+1} = p_m \sum_{k=1}^{\infty} p_k = p_m(1 - p_0).$$

Отже
$$p_m = (1 - p_0)^m p_0.$$
Це означає, що розподіл випадкової величини ξ геометричний.

Випадкова величина ξ має **розподіл Паскаля** з параметром $a > 0$, якщо вона має геометричний розподіл з параметром $p = 1/(1+a)$.

2.3.5 Негативний біноміальний розподіл

Дискретна випадкова величина ξ має негативний біноміальнй розподіл із параметрами p, r, позначення $\xi \simeq G(p, r)$, якщо вона збігається з кількістю випробувань до r-го успіху в нескінченній послідовності випробувань Бернуллі з імовірністю успіху p в одному випробуванні. Розподіл має вигляд
$$\mathbb{P}(\xi = k) = C_{k-1}^{r-1}(1-p)^{k-r}p^r,\ k = r, r+1, \ldots$$

Зауважимо, що кожна сприятлива елементарна подія довжини k містить $k - r$ неуспіхів та r успіхів, причому останнім має бути успіх. Тому ймовірність такої події дорівнює $(1-p)^{k-r}p^r$, а загальна кількість їх збігається з кількістю розміщень $r-1$ успіху на перших $k-1$ місць.

Вправа 2.1. Довести, що негативний біноміальний розподіл задовольняє таку рекурентну тотожність:
$$\mathbb{P}(\xi = k) = \mathbb{P}(\xi = k-1)(k-1)/(k-r).$$

2.3.6 Гіпергеометричний розподіл

Розглянемо такий стохастичний експеримент. В озері плаває всього N риб, із них n щук, а інші – карасі. Рибалка виловив m риб. Дискретна випадкова величина ξ, що дорівнює кількості щук у вилові, має гіпергеометричний розподіл із параметрами (N, n, m), позначення $\xi \simeq H(N, n, m)$.

Для обчислення розподілу величини ξ застосуэмо класичне означення ймовірностей та зауважимо, що кількість всіх елементарних подій дорівнює кількості способів обрати m риб із загального числа N, тобто C_N^m.

Кількість елементарних подій, що сприяють події $\{\xi = k\}$, дорівнює за основним правилом комбінаторики $C_n^k \cdot C_{N-n}^{m-k}$ – добуткові кількості способів обрати k щук із загальної кількості n та m˘k карасів із кількості $N-n$. Отже, гіпергеометричний розподіл має вигляд

$$\mathbb{P}(\xi = k) = \frac{C_n^k \cdot C_{N-n}^{m-k}}{C_N^m},\ k = 0, 1, \ldots, \min(n, m).$$

Зауваження 2.5. Якщо проінтерпретувати N як загальну кількість виробів, що виготовлені на підприємстві, n – як кількість бракованих серед них, m – як обсяг вибіркової партії, що була перевірена відділом контролю якості, то випадкова кількість виявлених бракованих виробів матиме гіпергеометричний розподіл. Цей факт використовують, зокрема, при статистичному контролі якості продукції.

Вправа 2.2. Показати, що найбільш імовірне значення k величини ξ з гіпергеометричним розподілом дорівнює $(n+1)(m+1)/(N+2)-1$.

Вказівка: відношення послідовних імовірностей має вигляд

$$(k+1)(N-n-m+k+1)/(n-k)(m-k).$$

2.3.7 Розподіл Пуассона

Як наступний приклад розглянемо випадкову величину ξ, яка набуває значення $0, 1, 2, \ldots, n, \ldots$ з імовірностями $p_0, p_1, p_2, \ldots, p_n, \ldots$, де

$$\mathbb{P}(\xi = k) = p_k = \frac{e^{-\lambda}\lambda^k}{k!} \quad k = 0, 1, 2, \ldots,$$

а $\lambda > 0$ – дійсна стала. Графік функції розподілу величини ξ подібний до сходів із нескінченною кількістю східців, зі стрибками в усіх цілих точках. Величина стрибка в точці $x = k$ дорівнює p_k. При $x \leq 0$ функція $F(x) = 0$. Про випадкову величину ξ кажуть, що вона розподілена за законом Пуассона з параметром λ (позначення $\xi \simeq Pois(\lambda)$).

Вправа 2.3. Довести, що розподіл Пуассона задовольняє таку рекурентну тотожність:

$$\mathbb{P}(\xi = k) = \mathbb{P}(\xi = k-1)\lambda/k.$$

Вправа 2.4. Знайти найбільш імовірне значення для розподілу Пуассона.

2.4 Абсолютно неперервні функції розподілу

Наступний клас містить виключно неперервні функції розподілу.

Означення 2.15. Випадкова величина ξ та її функція розподілу $F_\xi(x)$ називаються абсолютно неперервними, якщо існує функція $f_\xi(x)$ така, що

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(y)dy, \quad \forall x \in \mathbb{R}.$$

Інтеграл у цьому означенні слід розуміти як:

а) інтеграл Рімана – Стілтьєса для кусково-неперервної функції $f_\xi(x)$,

(б) інтеграл Лебега – Стілтьєса для вимірної функції $f_\xi(x)$.

Означення 2.16. Функція $f_\xi(x)$ називається щільністю розподілу випадкової величини ξ та функції розподілу $F_\xi(x)$.

Зауваження 2.6. Поняття інтеграла Лебега як часткового випадку математичного сподівання разом із його основними властивостями викладене нижче в розділі про загальне означення математичного сподівання.

З нормованості та монотонності функції розподілу випливають такі властивості щільності:

(1 –невід'ємність)

$$f_\xi(x) \geq 0, \quad \forall x \in \mathbb{R},$$

(2 –нормованість)

$$\int_{-\infty}^{\infty} f_\xi(y)dy = 1.$$

Будь-яка невід'ємна інтегровна нормована функція є щільністю певної функції розподілу, оскільки з властивостей інтегралу (Рімана або Лебега) випливають характеристичні властивості (1 – 3) функції розподілу.

Якщо функція розподілу диференційовна, то з формули Ньютона – Лейбніца (для інтеграла Рімана) або ж із теореми Радона – Нікодима (для інтеграла Лебега) випливає, що щільність (майже всюди) збігається з похідною функції розподілу

$$f_\xi(x) = F'_\xi(x) = dF_\xi(x)/dx.$$

В кожній точці неперервності x за теоремою про середнє має місце зображення

$$\mathbb{P}(x \le \xi < x+h) = \int_x^{x+h} f_\xi(y)dy = hf_\xi(x) + o(h), \ h \to 0,$$

тому $f_\xi(x)$ можна інтерпретувати як "щільність імовірності" в околі точки x.

Теорема 2.14. (Обчислення ймовірностей, пов'язаних із неперервною величиною.) *Якщо випадкова величина ξ абсолютно неперервна з щільністю $f_\xi(x)$), то*

$$\mathbb{P}(\xi \in B) = \int_B f_\xi(x)dx, \ \forall B \in \mathfrak{B}(\mathbb{R}).$$

Доведення. Шукана рівність виконується за означенням для напівінтервалів $B = (-\infty, x)$. З іншого боку, обидві частини рівності як функції B є сигма-адитивними мірами. Тому за теоремою Каратеодорі про продовження міри вони збігаються на породженій σ-алгебрі $\mathfrak{B}(\mathbb{R})$. □

Наслідок 2.2. (Щільність лінійного перетворення випадкової величини.) *Нехай випадкова величина ξ має щільність розподілу $f_\xi(x)$. Тоді лінійне перетворення $\zeta = a + b\xi$ має щільність*

$$f_\zeta(x) = \frac{1}{b} f_\xi((x-a)/b).$$

Доведення. Доведення виводиться з означення щільності та з формули заміни змінної в інтегралі $y = (u-a)/b$

$$\mathbb{P}(\zeta < x) = \mathbb{P}(\xi < (x-a)/b) =$$

$$= \int_{-\infty}^{(x-a)/b} f_\xi(y)dy = \frac{1}{b}\int_{-\infty}^{x} f_\xi((u-a)/b)du.$$

□

2.5 Сиргулярна функція розподілу

Крім дискретних і неперервних випадкових величин існують, звичайно, й інші випадкові величини. Крім величин, які поводять себе в одних інтервалах як неперервні, а в інших як дискретні, існують величини, які в жодному інтервалі не є ні дискретними, ні неперервними. До таких випадкових величин належать, наприклад, усі ті величини, функції розподілу яких неперервні, але зростають лише на множині міри нуль. Як приклад такої величини наведемо випадкову величину, що має функцією розподілу відому криву Кантора. Нагадаємо побудову цієї кривої. Випадкова величина набуває лише значення між нулем та одиницею. Отже, її функція розподілу задовольняє рівності $F(x) = 0$ при $x \leq 0$ та $F(x) = 1$ при $x > 1$. Усередині інтервалу $(0, 1)$ величина ξ набуває значення тільки в першій та останній третині, у кожній з імовірністю $1/2$. В інтервалі $\left(\frac{1}{3},\ \frac{2}{3}\right]$ функція $F(x)$ дорівнює $1/2$. В інтервалах $\left(0,\ \frac{1}{3}\right)$ та $\left(\frac{2}{3},\ 1\right)$ знову ξ може набувати значення лише в першій та останній третинах, у кожній з імовірністю $1/4$. Цим визначається $F(x)$ в інтервалах $\left(\frac{1}{9},\ \frac{2}{9}\right]$ та $\left(\frac{7}{9}, \frac{8}{9}\right]$: $F(x) = \frac{1}{4}$ при $\frac{1}{9} < x \leq \frac{2}{9}$, $\ F(x) = \frac{3}{4}$ при $\frac{7}{9} < x \leq \frac{8}{9}$. Далі в інтервалах $\left(0,\ \frac{1}{9}\right)$, $\left(\frac{2}{9},\ \frac{1}{3}\right)$, $\left(\frac{2}{3},\ \frac{7}{9}\right)$, $\left(\frac{8}{9},\ 1\right)$ повторюється знову та сама побудова, і ця побудова продовжується до нескінченності. У результаті $F(x)$ визначена в зліченній множині інтервалів, які є інтервалами деякої ніде не щільної досконалої множини. На цій множині функцію $F(x)$ визначаємо згідно з умовами: вона має бути неспадною та неперервною. Величина ξ не є дискретною, оскільки її функція розподілу неперервна, але в той же час вона не є неперервною, оскільки функція $F(x)$ не може бути виражена як інтеграл від щільності.

2.6 Класифікація функцій розподілу

Одночасно з класами абсолютно неперервних та дискретних функцій розподілу існує екзотичний клас сингулярних функцій розподілу, які є неперервними, але множина їх точок зростання має нульову довжину. Очевидно, що вказані три класи розподілів не перетинаються. Виявляється, що їх опукла оболонка вичерпує весь клас функцій розподілу.

Теорема 2.15. (Теорема Лебега про зображення функції розподілу.) *Для довільної функції розподілу $F(x)$ існують: дискретна $F_d(x)$, абсолютно неперервна $F_a(x)$ і сингулярна $F_s(x)$ фун-*

кції розподілу та три числа $\alpha \geq 0, \beta \geq 0, \gamma \geq 0$, *такі, що* $\alpha + \beta + \gamma = 1$ *та*

$$F(x) = \alpha F_d(x) + \beta F_a(x) + \gamma F_s(x),\ \forall x \in \mathbb{R}.$$

2.7 Неперервні розподіли

Розглянемо приклади найбільш вживаних випадкових величин які мають неперервні функції розподілу.

2.7.1 Рівномірний розподіл

Означення 2.17. Випадкова величина ξ має рівномірний розподіл на відрізку $[a, b]$, позначення $\xi \simeq U(a, b)$, якщо її щільність є сталою всередині цього відрізку, та дорівнює нулю поза ним.

Отже, ймовірність потрапляння величини у множину всередині відрізку як інтеграл від щільності пропорційна довжині цієї множини і не залежить від її положення – тобто виконується умова рівноймовірності значень. З умови нормованості виводимо, що щільність розподілу рівномірної на $[a, b]$ випадкової величини дорівнює

$$f(x) = \begin{cases} 0, & x \notin [a, b], \\ \frac{1}{b-a}, & x \in [a, b], \end{cases}$$

а функція розподілу має вигляд

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

2.7.2 Нормальний (Гауссів) розподіл

Означення 2.18. Випадкова величина ξ має стандартний нормальний розподіл, позначення $\xi \simeq N(0, 1)$, якщо її щільність дорівнює

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

тобто функція розподілу має вигляд

$$F_\xi(x) = \int\limits_{-\infty}^{x} \varphi(y) dy = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}} dy.$$

Часто цей розподіл називають також гауссовим розподілом.

Нормованість, $F(+\infty) = 1$, є наслідком виразу для інтеграла Пуассона

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

Означення 2.19. Випадкова величина ξ має нормальний розподіл із параметрами a та $\sigma > 0$, позначається $\xi \simeq N(a, \sigma^2)$, якщо щільність розподілу величини ξ має вигляд

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\},$$

Випадкову величину ξ можна записати у вигляді лінійного перетворення $\xi = a + \sigma\zeta$, де ζ – стандартна нормальна величина, тобто має стандартний нормальний розподіл.

Далі ми виявимо ймовірнісний сенс величин a та σ.

Функція $\varphi(x)$ досягає максимуму при $x = a$ і має точки перегину при $x = a \pm \sigma$; вісь абсцис служить для неї асимптотою при $x \to \pm\infty$. Ми бачимо, що чим менше значення σ, тим крива $\varphi(x)$ має більше значення максимуму і спадає крутіше. Це означає, зокрема, що ймовірність потрапити в інтервал $(-b, b)$ більша для тієї випадкової величини, розподіленої за нормальним законом (з параметром $a = 0$), для якої величина σ менша. Отже, ми можемо вважати σ характеристикою розсіювання значень величини ξ. При $a \neq 0$ криві щільності мають ту ж форму, але зсунуті вправо ($a > 0$) або вліво ($a < 0$) залежно від знака параметра а.

Нормально розподілені величини відіграють особливо важливу роль у теорії ймовірностей та її застосуваннях, далі будемо мати багато можливостей переконатися в цьому.

2.7.3 Логнормальний розподіл

Означення 2.20. Випадкова величина ξ має логнормальний розподіл, позначення $\xi \simeq LN(a, \sigma^2)$, якщо її логарифм має нормальний розподіл з відповідними параметрами: $\ln \xi \simeq N(a, \sigma^2)$.

Випадкову величину ξ можна зобразити у вигляді $\xi = \exp(\zeta)$, де $\zeta \simeq N(a, \sigma^2)$.

Нормальний розподіл є граничним для сум великого числа незалежних величин, тоді як логнормальний розподіл є граничним у мультиплікативних схемах, що містять добутки незалежних факторів.

2.7.4 Показниковий (експоненційний) розподіл

Означення 2.21. Випадкова невід'ємна величина ξ має показниковий (експоненційний) розподіл з параметром $\lambda > 0$, позначення $\xi \simeq Exp(\lambda)$), якщо її щільність розподілу має вигляд

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

Показниково розподілена випадкова величина має гама розподіл з параметрами $\alpha = 1, \beta = \lambda$, тобто $\xi \simeq Exp(\lambda) \simeq \Gamma(1, \lambda)$.

Теорема 2.16. (Про відсутність післядії для показникового розподілу.) *Показниково розподілена випадкова величина ξ має властивість відсутності післядії, тобто для довільних $t, s \geq 0$ має місце тотожність*

$$\mathbb{P}(\xi \geq t + s | \xi \geq t) = \mathbb{P}(\xi \geq s).$$

Цю властивість можна проінтерпретувати так: за умови "виживання" показникова величина повністю забуває своє минуле.

Доведення. Доведення теореми є очевидним наслідком мультиплікативності експоненти $\exp(-\lambda x)$ та означення умовної ймовірності:

$$\mathbb{P}(\xi \geq t + s | \xi \geq t) = \frac{\mathbb{P}(\xi \geq t + s, \xi \geq t)}{\mathbb{P}(\xi \geq t)} =$$

$$= \frac{\mathbb{P}(\xi \geq t + s)}{\mathbb{P}(\xi \geq t)} = \frac{\exp(-\lambda(t + s))}{\exp(-\lambda t)} = \mathbb{P}(\xi \geq s).$$

□

Вправа 2.5. Довести, що в класі невід'ємних та необмежених величин властивість відсутності післядії мають лише показникові величини. Вказівка: відсутність післядії еквівалентна мультиплікативності $Q(t + s) = Q(t)Q(s)$ для ймовірності виживання $Q(t) = \mathbb{P}(\xi \geq t)$, а єдиним монотонним розв'язком рівняння $Q(t + s) = Q(t)Q(s)$ з $Q(0) = 1$ є експонента.

2.7.5 Гама розподіл

Випадкова невід'ємна величина ξ має гама розподіл з параметрами (α, β) (позначення $\xi \simeq \Gamma(\alpha, \beta)$), якщо її щільність розподілу має вигляд

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0. \end{cases}$$

Властивості гама розподілу базуються на властивостях гама-функції

$$\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx.$$

Найбільш поширені власимвості

$$\Gamma(n+1) = n!,\ \Gamma(n+1) = n\Gamma(n),\ \Gamma(1) = 1,$$

$$\Gamma(1/2) = \sqrt{\pi},\ \Gamma(\alpha+1) = \alpha\Gamma(\alpha).$$

2.7.6 Розподіл Ерланга

Випадкова невід'ємна величина ξ має розподіл Ерланга з параметрами n, λ (позначення $\xi \simeq Erlang(n, \lambda)$), якщо її щільність розподілу має вигляд

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, & x \geq 0. \end{cases}$$

Випадкова величина яка має розподіл Ерланга з параметрами n, λ має гама розподіл з параметрами $\alpha = n, \beta = \lambda$, тобто $\xi \simeq Erlang(n, \lambda) \simeq \Gamma(n, \lambda)$.

2.7.7 Розподіл хі-квадрат

Випадкова невід'ємна величина ξ має розподіл хі-квадрат з параметром n (позначення $\xi \simeq \chi_n^2$), якщо її щільність розподілу має вигляд

$$p(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}.$$

Така випадкова величина має гама розподіл з параметрами $\alpha = \frac{n}{2}n, \beta = \frac{1}{2}$, тобто $\xi \simeq \chi_n^2 \simeq \Gamma(\frac{n}{2}, \frac{1}{2})$.

2.7.8 Функція інтенсивності розподілу

Нехай випадкова величина ξ має щільність розподілу $f(x)$ і функцію розподілу $F(x)$. Значення $f(x)$ (за припущенням неперервності $f(x)$) можна інтерпретувати як границю

$$f(x) = \lim_{h\to 0} \frac{\mathbb{P}(x \le \xi < x+h)}{h} = \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(y)dy.$$

На практиці одночасно з безумовною ймовірністю

$$\mathbb{P}(x \le \xi < x+h)$$

важливе значення відіграє умовна ймовірність

$$\mathbb{P}(x \le \xi < x+h | \, \xi \ge x).$$

Саме вона вказує на частку відмов за одиницю часу серед тих приладів, які не відмовили на початок поточного періоду, в той час як безумовна ймовірність задає частку відмов серед усіх приладів, що спостерігалися з самого початку випробувань. Відповідним умовним аналогом щільності є функція, що дорівнює границі для умовних імовірностей

$$\lambda(x) = \lim_{h\to 0} \frac{\mathbb{P}(x \le \xi < x+h | \, \xi \ge x)}{h}$$

За означенням умовної ймовірності

$$\mathbb{P}(x \le \xi < x+h | \, \xi \ge x) = \frac{\mathbb{P}(x \le \xi < x+h, \xi \ge x)}{\mathbb{P}(\xi \ge x)} =$$

$$= \frac{\mathbb{P}(x \le \xi < x+h)}{\mathbb{P}(\xi \ge x)} = h\frac{f(x)}{(1-F(x))} + o(h), \; h \to 0.$$

Звідси приходимо до такого означення.

Означення 2.22. Функцією інтенсивності невід'ємної випадкової величини ξ, що має щільність $f(x)$ і функцію розподілу $F(x)$, називається функція

$$\lambda(x) = \frac{f(x)}{(1-F(x))}, \, x \ge 0.$$

За означенням для всіх точок неперервності x щільності f

$$\mathbb{P}(x \le \xi < x+h|\xi \ge x) = h\lambda(x) + o(h), \, h \to 0.$$

Теорема 2.17. (Про обчислення розподілу через функцію інтенсивності.) *Нехай випадкова величина ξ невід'ємна, абсолютно неперервна і має функцію інтенсивності $\lambda(x)$. Тоді її функція розподілу та щільність мають при $x \geq 0$ вигляд*

$$F(x) = 1 - \exp\left(-\int_0^x \lambda(y)dy\right), \quad f(x) = \lambda(x)\exp\left(-\int_0^x \lambda(y)dy\right).$$

Доведення. Обчислимо інтеграл

$$\int_0^x \lambda(y)dy = \int_0^x \frac{f(y)}{(1-F(y))}dy = \int_0^x \frac{1}{(1-F(y))}dF(y) =$$

$$\int_{F(0)}^{F(x)} \frac{1}{(1-u)}du = -\ln(1-F(x))$$

із заміною змінної $u = F(y)$ та урахуванням того, що щільність $f(x)$ є (майже скрізь) похідною для функції розподілу $F(x)$. Розв'язком цього рівняння є наведена рівність для $F(x)$. Із монотонності та нормованості функції розподілу виводимо характеристичні властивості інтенсивності

$$\lambda(x) \geq 0, \quad \int_0^\infty \lambda(x)dx = \infty.$$

Далі, заміною змінної

$$u = \int_0^x \lambda(y)dy$$

обчислимо інтеграл

$$\int_0^t \lambda(x)\exp\left(-\int_0^x \lambda(y)dy\right)dt = 1 - \exp\left(-\int_0^t \lambda(y)dy\right) = F(t),$$

звідки за означенням щільності дістанемо зображення для щільності. □

Показниково розподілена випадкова величина має сталу функцію інтенсивності, що дорівнює λ.

2.7.9 Розподіл Вейбула

Важливими є розподіли зі змінною інтенсивністю.

Означення 2.23. Випадкова величина ξ має розподіл Вейбула з параметрами $\lambda > 0$, $a > 0$, якщо її щільність розподілу має вигляд

$$f(x) = \begin{cases} 0, & x < 0, \\ \lambda a(\lambda x)^{a-1}\exp\left(-(\lambda x)^{a}\right), & x \geq 0. \end{cases}$$

Функція розподілу Вейбула використовується в теорії стохастичних моделей як більш реальний замінник показникової функції розподілу, оскільки властивість відсутності післядії у більшості застосувань не є прийнятною. Дійсно, більш реалістичний характер поведінки функції інтенсивності полягає в тому, що на початковому етапі вона спадає (ефект початкових конструктивних відмов), лише потім наступає період відносної сталості, що змінюється етапом неухильного зростання (ефект старіння).

2.7.10 Розподіл Гомпертца

Означення 2.24. Випадкова величина ξ має розподіл Гомпертца з параметрами $\lambda > 0, \mu > 0, a > 0$, якщо її функція інтенсивності є показниковою

$$\lambda(x) = \lambda + \mu\exp(ax),\ x \geq 0.$$

Розподіл Гомпертца використовується у демографії як математична модель тривалості життя людини.

2.7.11 Бета-розподіл

Означення 2.25. Випадкова величина ξ має бета-розподіл на $(0, 1)$ з параметрами $\alpha > 0, \beta > o$, якщо її щільність розподілу має вигляд:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)},\ x \in (0,1),$$

де $B(\alpha, \beta)$ – повна бета-функція.

При $\alpha = 1, \beta = 1$ даний розподіл збігається з рівномірним розподілом. У інших ипадках він відбиває можливу помірну варіативність щільності.

2.7.12 Розподіл Парето

Зазначені вище функції розподілу прямують до одиниці при $x \to \infty$ з експоненційною швидкістю. Для розподілу Парето ця швидкість є степеневою, що є суттєвим у деяких моделях.

Означення 2.26. Випадкова величина ξ має розподіл Парето, якщо для деяких $\lambda > 0, \alpha > 0$ її функція розподілу при $x \geq 1/\lambda$ дорівнює

$$\mathbb{P}(\xi < x) = 1 - (\lambda x)^{-\alpha}.$$

2.7.13 Розподіл Коші

Означення 2.27. Випадкова величина ξ має розподіл Коші, якщо її можна зобразити у вигляді $\xi = \mathrm{tg}(\varphi)$, де випадковий кут φ є рівномірно розподіленим на відрізку $(-\pi/2, \pi/2)$. Оскільки $\varphi \in (-\pi/2, \pi/2)$, то

$$\mathbb{P}(\xi < x) = \mathbb{P}(\varphi < \mathrm{arctg}(x)) = \frac{1}{2} + \frac{1}{\pi}\mathrm{arctg}(x),\ x \in \mathbb{R}$$

і щільність ξ має вигляд

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2},\ x \in \mathbb{R}.$$

2.8 Багатовимірні функції розподілу

Якщо кожній елементарній події ω поставити у відповідність не одне дійсне число, а сукупність k дійсних чисел $\xi_1, \xi_2, \ldots, \xi_k$, то, припускаючи вимірність множин типу

$$\{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_k < x_k\}$$

при довільних дійсних $x_1, x_2, \ldots, x_k$ відносно визначеної у просторі $(\Omega, \mathfrak{F}, \mathbb{R})$ σ-алгебри випадкових подій, назвемо систему з k величин $(\xi_1, \xi_2, \ldots, \xi_k)$ випадковим вектором, або k -вимірною випадковою величиною.

Імовірність здійснення нерівностей

$$\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_k < x_k$$

за всіх можливих значень $x_1, x_2, \ldots, x_k$, тобто функція

$$F\left(x_1, x_2, \ldots, x_k\right) = P\left\{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_k < x_k\right\}$$

називається функцією розподілу k -вимірної випадкової величини.

Для геометричної ілюстрації зображатимемо величини $\xi_1, \xi_2, \ldots, \xi_k$ як координати точок k-вимірного евклідового простору. Очевидно, що положення точки $(\xi_1, \ldots, \xi_k)$ у просторі залежить від випадку і що функція $F(x_1, \ldots, x_k)$ за такої інтерпретації дає ймовірність потрапляння точки $(\xi_1, \ldots, \xi_k)$ в k-вимірний паралелепіпед

$$\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_k < x_k$$

з ребрами, паралельними до осей координат. За допомогою функцій розподілу легко обчислити ймовірність того, що точка $(\xi_1, \ldots, \xi_k)$ міститиметься всередині паралелепіпеда

$$a_i \le \xi_i < b_i, \quad i = 1, 2, \ldots, k,$$

де a_i та b_i – довільні дійсні числа. Неважко підрахувати, що

$$P\{a_1 \le \xi_1 < b_1,\ a_2 \le \xi_2 < b_2, \ldots, a_k \le \xi_k < b_k\} =$$

$$= F(b_1, b_2, \ldots, b_k) - \sum_{i=1}^{k} p_i + \sum_{i<j} p_{ij} - \ldots + (-1)^k F(a_1, a_2, \ldots, a_k), \tag{2.1}$$

де $p_{ij\ldots s}$ – значення функції $F(c_1, c_2, \ldots, c_k)$ при $c_i = a_i$, $c_j = a_j$, $\ldots$, $c_s = a_s$ і при інших c_r, рівних b_r.

Зауважимо, що

$$F(x_1, \ldots, x_{s-1}, +\infty, x_{s+1}, \ldots, x_k)$$

дає нам імовірність того, що буде справедливою така система нерівностей

$$\xi_1 < x_1,\ \xi_2 < x_2,\ \ldots,\ \xi_{s-1} < x_{s-1},\ \xi_{s+1} < x_{s+1}, \ldots, \xi_k < x_k\,.$$

Оскільки за аксіомою додавання ймовірностей

$$P\{\xi_1 < x_1, \ldots, \xi_{s-1} < x_{s-1}, \xi_{s+1} < x_{s+1}, \ldots, \xi_k < x_k\} =$$
$$\sum_{r=-\infty}^{\infty} P\{\xi_1 < x_1, \ldots, \xi_{s-1} < x_{s-1}, r \le \xi_s < r+1, \xi_{s+1} < x_{s+1}, \ldots, \xi_k < x_k\} =$$
$$= F(x_1, \ldots, x_{s-1}, +\infty, x_{s+1}, \ldots, x_k)\,,$$

то

$$F(x_1, \ldots, x_{s-1}, +\infty, x_{s+1}, \ldots, x_k)$$

є функцією розподілу $(k-1)$-вимірної випадкової величини $(\xi_1, \ldots, \xi_{s-1}, \xi_{s+1}, \ldots, \xi_k)$. Продовжуючи цей процес далі, можемо визначити r -вимірні функції розподілу довільної групи з r величин $\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_r}$ за формулою

$$F\left(x_{i_1}, x_{i_2}, \ldots, x_{i_r}\right) = P\left\{\xi_{i_1} < x_{i_1}, \ldots, \xi_{i_r} < x_{i_r}\right\} = F\left(c_1,\, c_2, \ldots, c_k\right), \tag{2.2}$$

де $c_s = x_s$, якщо $s = i_j$ $(1 \le j \le r)$ і $c_s = +\infty$ в інших випадках.

Зокрема, функція розподілу випадкової величини ξ_i дорівнює

$$F_i\left(x\right) = P\left\{\xi_i < x\right\} = F\left(c_1, c_2, \ldots, c_k\right),$$

де всі c_j $(j \ne i)$ дорівнюють $+\infty$, а $c_i = x$.

Розглянемо приклади k -вимірних функцій розподілу.

2.8.1 Рівномірний розподіл у паралелепіпеді

Випадковий вектор $(\xi_1,\, \xi_2, \ldots, \xi_k)$ називається рівномірно розподіленим у паралелепіпеді $a_i \le \xi_i < b_i$, $1 \le i \le k$, якщо ймовірність потрапляння точки $(\xi_1,\, \xi_2, \ldots, \xi_k)$ в довільну внутрішню область указаного паралелепіпеда пропорційна її об'єму, а потрапляння всередину паралелепіпеда є достовірною подією. Функція розподілу рівномірно розподіленої в паралелепіпеді $a_i \le \xi_i < b_i$, $1 \le i \le k$ випадкової величини має вигляд

$$F\left(x_1, x_2, \ldots, x_k\right) = \begin{cases} 0, & \text{якщо } x_i \le a_i \text{ хоча б при одному } i; \\ \prod\limits_{i=1}^{k} \frac{c_i - a_i}{b_i - a_i}, & c_i = \begin{cases} x_i, \text{ якщо } a_i \le x_i \le b_i, \\ b_i, \text{ якщо } x_i > b_i. \end{cases} \end{cases}$$

2.8.2 Двовимірний нормальний розподіл

Двовимірна випадкова величина (ξ_1, ξ_2) розподілена за нормальним законом, якщо для неї функція розподілу дорівнює

$$F\left(x_1, x_2\right) = C \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{x_2} e^{-Q(x,\, y)} dy\, dx,$$

де $Q\left(x,\, y\right)$ – додатно визначена квадратична форма. Відомо, що додатно визначена квадратична форма від двох змінних x та y може бути записана у вигляді

$$Q\left(x,\, y\right) = \frac{1}{2}\left\{\frac{\left(x-a\right)^2}{A^2} - 2r\frac{\left(x-a\right)\left(y-b\right)}{AB} + \frac{\left(y-b\right)^2}{B^2}\right\},$$

де A та B – додатні числа, a, b, r – дійсні числа; причому r задовольняє умову $-1 \leq r \leq 1$. Легко бачити, що при $r^2 \neq 1$ кожна з випадкових величин ξ_1, ξ_2 має одновимірний нормальний розподіл. Справді,

$$F_1(x_1) = P\{\xi_1 < x_1\} = F(x_1, +\infty) = C \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} e^{-Q(x,\, y)}\, dy\, dx =$$

$$= C \int_{-\infty}^{x_1} e^{-\frac{(x-a)^2}{2A^2}\left(1-r^2\right)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{y-b}{B} - r\frac{x-a}{A}\right]^2} dy dx.$$

Оскільки

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{y-b}{B} - r\frac{x-a}{A}\right]^2} dy \ = B\sqrt{2\pi},$$

то

$$F_1(x_1) = BC\sqrt{2\pi} \int_{-\infty}^{x_1} e^{-\frac{(x-a)^2}{2A^2}\left(1-r^2\right)} dx. \tag{2.3}$$

Стала C може бути виражена через A, B та r. Цю залежність можна знайти з умови $F_1(+\infty) = 1$. Маємо

$$1 = BC\sqrt{2\pi} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2A^2}\left(1-r^2\right)} dx = \frac{ABC\sqrt{2\pi}}{\sqrt{1-r^2}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \frac{ABC(2\pi)}{\sqrt{1-r^2}}.$$

Звідси

$$C = \frac{\sqrt{1-r^2}}{2\pi AB}.$$

Якщо $r^2 \neq 1$, то покладемо

$$A = \sigma_1\sqrt{1-r^2}, \quad B = \sigma_2\sqrt{1-r^2}.$$

У цих нових позначеннях двомірний нормальний закон набуває такого вигляду:

$$F(x_1,\, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} e^{-\frac{1}{(1-r^2)}\left(\frac{(x-a)^2}{2\sigma_1^2} - r\frac{(x-a)(y-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{2\sigma_2^2}\right)} dy\, dx.$$

Теоретико-ймовірнісний сенс сталих a, b, r, σ_1 та σ_2 буде з'ясовано в наступному розділі.

При $r^2 = 1$ рівність (2.3) утрачає сенс. Далі побачимо, що в цьому випадку величини ξ_1 та ξ_2 зв'язані між собою лінійною залежністю.

2.8.3 Властивості багатовимірної функції розподілу

Подібно до того, як ми це робили в одновимірному випадку, для багатовимірних функцій розподілу можна встановити низку властивостей.

Функція розподілу $F(x_1, x_2, \ldots, x_k)$

1) є неспадна функція кожного аргументу,

2) неперервна зліва за кожним аргументом,

3) задовольняє співвідношення

$$F(+\infty, +\infty, \ldots, +\infty) = 1,$$

$$\lim_{x_i \to -\infty} F(x_1, x_2, \ldots, x_k) = 0, \quad 1 \le i \le k,$$

за довільних значень інших аргументів.

В одновимірному випадку було видно, що перелічені властивості необхідні й достатні, щоб функція $F(x)$ була функцією розподілу деякої величини; у багатовимірному випадку цих властивостей вже недостатньо.

Щоб функція $F(x_1, x_2, \ldots, x_k)$ була функцією розподілу, необхідно, крім попередніх трьох властивостей, додати ще таку:

4) за довільних a_i та b_i $(i = 1, 2, \ldots, k)$ вираз (2.1) невід'ємний.

Те, що ця вимога може бути не виконана, незважаючи на наявність у функції $F(x_1, x_2, \ldots, x_k)$ властивостей 1) – 3), показує такий приклад.

Нехай

$$F(x, y) = \begin{cases} 0, & x \le 0, \quad x + y \le 1, \quad y \le 0, \\ 1 & \text{для інших}\, x, \; y. \end{cases}$$

Ця функція задовольняє вимоги 1) – 3), але для неї

$$F(1, 1) - F\left(1, \frac{1}{2}\right) - F\left(\frac{1}{2}, 1\right) + F\left(\frac{1}{2}, \frac{1}{2}\right) = -1 \qquad (2.4)$$

і, отже, четверта вимога не виконується.

Функція $F(x, y)$ не може бути функцією розподілу, оскільки різниця (2.4) згідно із співвідношенням (2.1) дорівнює ймовірності потрапляння точки (ξ_1, ξ_2) у прямокутник

$$\left(\frac{1}{2} \le x < 1, \frac{1}{2} \le y < 1\right).$$

2.8.4 Сумісна щільність розподілу випадкового вектора

Якщо існує така невід'ємна вимірна функція $p\,(x_1,\,x_2,\ldots,x_k)$, що за довільних $x_1,\,x_2,\ldots,x_k$ має місце рівність

$$F\,(x_1,\,x_2,\ldots,\,x_k)=\int\limits_{-\infty}^{x_1}\int\limits_{-\infty}^{x_2}\ldots\int\limits_{-\infty}^{x_k}p\,(z_1,\,z_2,\ldots,z_k)\;dz_k\ldots dz_1,$$

то ця функція називається щільністю розподілу ймовірностей випадкового вектора $(\xi_1,\,\xi_2,\ldots,\xi_k)$.

З нормованості та монотонності функції розподілу випливають такі властивості щільності розподілу:

(1 - невід'ємність) $p\,(x_1,\,x_2,\,\ldots\,,\,x_k)\geq 0$,

(2 - нормованість)

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty}p\,(z_1,z_2,\ldots,z_k)\;dz_k\ldots dz_1=1.$$

Будь-яка невід'ємна вимірна нормована функція є сумісною щільні стю певної сумісної функції розподілу, оскільки з властивостей інтеграла (Рімана або Лебега) випливатимуть характеристичні властивості (1)-(4) сумісної функції розподілу, а отже, може бути побудована відповідна мі ра Лебега – Стілтьєса, що породжена сумісною функцією розподілу.

Якщо сумісна функція розподілу диференційовна, то з формули Ньютона – Лейбніца (для інтеграла Рімана) або ж із теореми Радона – Нікодима (для інтеграла Лебега) випливає, що сумісна щільність (майже всюди) збігається з похідною функції розподілу:

$$p_\xi(x_1,x_2,\ldots,x_k)=\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\cdots\frac{\partial}{\partial x_k}F_\xi(x_1,x_2,\ldots,x_k).$$

Теорема 2.18. (Про обчислення ймовірностей, пов'язаних із неперервним вектором.)

Якщо випадковий вектор $(\xi_1,\xi_2,\ldots,\xi_k)$ має сумісну щільність $p_\xi(x_1,x_2,\ldots,x_k)$, то для всіх борелевих $B\in\mathfrak{B}(\mathbb{R}^k)$

$$\mathbb{P}(\xi\in B)=\int\underset{B}{\ldots}\int p\,(x_1,x_2,\ldots,x_k)\;dx_1\ldots dx_k.$$

Доведення. Шукана рівність виконується для кутів $B = (-\infty, x)$ за означенням. З іншого боку, обидві частини рівності як функції множини B є сигма-адитивними мірами. Тому за теоремою Каратеодорі про продовження міри вони збігаються на породженій σ-алгебрі $B \in \mathfrak{B}(\mathbb{R}^k)$. □

Зауважимо, що якщо $p\left(x_1, x_2, \ldots, x_k\right)$ неперервна в точці $(x_1, \ldots, x_k)$, то ймовірність вектора $(\xi_1, \xi_2, \ldots, \xi_k)$ потрапити всередину паралелепіпеда

$$x_j \leq \xi_j < x_j + dx_j \ \ (j = 1, \ldots, k)$$

з точністю до нескінченно малих вищих порядків дорівнює

$$p\ (x_1, x_2, \ldots, x_k)\ dx_1, dx_2, \ldots, dx_k.$$

2.8.5 Рівномірний розподіл в області

Як приклад k -вимірної випадкової величини, яка має щільність, наведемо випадкову величину, рівномірно розподілену в k -вимірній області G . Якщо через v позначимо k -вимірний об'єм області G, то щільність розподілу дорівнюватиме

$$p\ (x_1, x_2, \ldots, x_k) = \begin{cases} 0, & (x_1,\ x_2, \ldots, x_k) \notin G, \\ \frac{1}{v}, & (x_1, x_2, \ldots, x_k) \in G. \end{cases}$$

2.8.6 Щільність двовимірного нормального закону

Щільність розподілу двовимірного нормального закону визначається формулою

$$p\ (x_1,\ x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x_1-a)^2}{\sigma_1^2} - 2r\frac{(x_1-a)(x_2-b)}{\sigma_1\sigma_2} + \frac{(x_2-b)^2}{\sigma_2^2}\right]}.$$

Зауважимо, що щільність розподілу нормального закону зберігає стале значення на еліпсах

$$\frac{\left(x_1-a\right)^2}{\sigma_1^2} - 2r\frac{\left(x_1-a\right)\left(x_2-b\right)}{\sigma_1\sigma_2} + \frac{\left(x_2-b\right)^2}{\sigma_2^2} = \lambda^2, \qquad (2.5)$$

де λ – стала; на цій підставі еліпси (2.5) носять назву еліпсів рівної ймовірності. Знайдемо ймовірність потрапляння точки $(\xi_1,\ \xi_2)$

всередину еліпса (2.5). Ця ймовірність, за властивістю щільності, дорівнює

$$P(\lambda) = \iint\limits_{G(\lambda)} p(x,\, y)\; dy\, dx, \tag{2.6}$$

де через $G(\lambda)$ позначено область, що обмежена еліпсом (2.5). Для обчислення цього інтеграла введемо полярні координати

$$x - a = \rho\ \cos\theta, \quad y - b = \rho\ \sin\theta.$$

Інтеграл (2.6) при цьому набуває вигляду

$$P(\lambda) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int\limits_0^{2\pi\lambda/s} \int\limits_0^{\sqrt{1-r^2}} e^{-\frac{\rho^2}{2}s^2} \rho\, d\rho\, d\theta,$$

де для скорочення позначено

$$s^2 = \frac{1}{1-r^2}\left[\frac{\cos^2\theta}{\sigma_1^2} - 2r\frac{\cos\theta\ \sin\theta}{\sigma_1\sigma_2} + \frac{\sin^2\theta}{\sigma_2^2}\right].$$

Інтегрування за ρ дає

$$P(\lambda) = \frac{1 - e^{-\frac{\lambda^2}{2(1-r^2)}}}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int\limits_0^{2\pi} \frac{d\theta}{s^2}.$$

Інтегрування за θ можна виконати за правилами інтегрування тригонометричних функцій, проте простіше скористатися рівністю

$$P(+\infty) = 1 = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int\limits_0^{2\pi} \frac{d\theta}{s^2}.$$

Звідси

$$\int\limits_0^{2\pi} \frac{d\theta}{s^2} = 2\pi\sigma_1\sigma_2\sqrt{1-r^2},$$

і, отже,

$$P(\lambda) = 1 - e^{-\frac{\lambda^2}{2(1-r^2)}}.$$

Нормальний розподіл відіграє винятково важливу роль у різних прикладних задачах. Розподіл багатьох практично важливих випадкових величин виявляється підпорядкованим нормальному закону. Так, практика проведення стрільби за різних умов показала, що розсіювання снарядів на площині при стрільбі з однієї гармати на певному прицілі підпорядковано нормальному закону. Ознайомившись з теоремою Ляпунова, побачимо, що ця універсальність нормального закону пояснюється тим, що випадкова величина, яка є сумою дуже великої кількості незалежних випадкових величин, кожна з яких має лише незначний вплив на суму, розподілена за майже нормальним законом. Щоб не створювати у читача помилкового уявлення, слід зараз же відзначити, що нормальний закон не є єдиним, абсолютно універсальним: багато реальних явищ підлягають іншим імовірнісним закономірностям, у чому зможемо згодом неодноразово переконатися.

2.9 Незалежні випадкові величини

Найважливіше поняття теорії ймовірностей – незалежність подій зберігає своє значення і для випадкових величин. Згідно з визначенням незалежності подій скажемо, що випадкові величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні, якщо для довільної групи $\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_k}$ цих величин має місце рівність

$$P\left\{ \xi_{i_1} < x_{i_1}, \xi_{i_2} < x_{i_2}, \ldots, \xi_{i_k} < x_{i_k} \right\} =$$
$$= P\left\{ \xi_{i_1} < x_{i_1} \right\} \cdot P\left\{ \xi_{i_2} < x_{i_2} \right\} \ldots P\left\{ \xi_{i_k} < x_{i_k} \right\} \qquad (2.7)$$

при довільних дійсних числах $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ і довільному k $(1 \leq k \leq n)$. Зокрема, для довільних $x_1, x_2, \ldots, x_n$ виконується рівність

$$P\left\{ \xi_1 < x_1, , \xi_2 < x_2, \ldots, \xi_n < x_n \right\} = \prod_{i=1}^{n} P\left\{ \xi_i < x_i \right\},$$

або в термінах функцій розподілу

$$F\left(x_1,\, x_2, \ldots, x_n\right) = \prod_{i=1}^{n} F_i\left(x_i\right),$$

де $F_i(x_i)$ означає функцію розподілу випадкової величини ξ_i.

Важливо зауважити, що має місце й обернене твердження: якщо функція розподілу $F(x_1, x_2, \ldots, x_n)$ системи випадкових величин ξ_1, ξ_2, ..., ξ_n має вигляд

$$F(x_1,\, x_2,\, \ldots\, ,\, x_n) = \prod_{i=1}^{n} F_i(x_i), \tag{2.8}$$

де функції $F_i(x_i)$ задовольняють співвідношення

$$F_i(+\infty) = 1 \ (1 \le i \le n)\,,$$

то величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні й

$$F_1(x_1)\,, F_2(x_2)\,, \ldots, F_n(x_n)$$

– є їхніми функціями розподілу. Дійсно, із сформульованих раніше властивостей 1) – 3) n-вимірних функцій розподілу випливає, що функції $F_i(x_i)$ неспадні, неперервні зліва і задовольняють вимоги $F_i(-\infty) = 0$, $F_i(+\infty) = 1$, тобто мають властивості функцій розподілу одновимірних випадкових величин.

Ми бачили раніше, що функція розподілу величини ξ_i визначається за формулою

$$P\{\,\xi_i < x_i\} = F(c_1, c_2, \ldots, c_n)\,,$$

де $c_k = +\infty$ при $k \ne i$ і $c_k = x_i$ при $k = i$. Таким чином, з (2.8) знаходимо, що

$$P\{\,\xi_i < x_i\} = F_i(x_i)\,.$$

Так само рівності (2.2) та (2.8) дозволяють для довільної групи $\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_k}$ випадкових величин визначити її функцію розподілу

$$P\{\,\xi_{i_1} < x_{i_1}, \xi_{i_2} < x_{i_2}, \ldots, \xi_{i_k} < x_{i_k}\} = \prod_{j=1}^{k} F_{i_j}\left(x_{i_j}\right).$$

Приклад 2.1. Розглянемо n-вимірну випадкову величину $(\xi_1, \xi_2, ..., \xi_n)$, компоненти $\xi_1, \xi_2, ..., \xi_n$ якої взаємно незалежні випадкові величини, розподілені за нормальним законом

$$F_k(x_k) = \frac{1}{\sigma_k\sqrt{2\pi}} \int\limits_{-\infty}^{x_k} e^{-\frac{(z-a_k)^2}{2\sigma_k^2}}\, dz.$$

У цьому прикладі функція розподілу дорівнює

$$F\left(x_1,\, x_2, \ldots, x_n\right) = \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{k=1}^{n} \frac{1}{\sigma_k} \int_{-\infty}^{x_k} e^{-\frac{(z-a_k)^2}{2\sigma_k^2}} dz.$$

Якщо незалежні випадкові величини $\xi_1, \xi_2, ..., \xi_n$ мають щільності $p_1\left(x_1\right), p_2\left(x_2\right), \ldots, p_n\left(x_n\right)$, то n-вимірна величина $\left(\xi_1, \xi_2, \ldots, \xi_n\right)$ має щільність розподілу, рівну

$$p\left(x_1, x_2, \ldots, x_n\right) = p_1\left(x_1\right) \cdot p_2\left(x_2\right) \cdots p_n\left(x_n\right). \tag{2.9}$$

Зі сказаного вище випливає, що коли щільність n-вимірної випадкової величини $\left(\xi_1, \xi_2, \ldots, \xi_n\right)$ може бути виражена у вигляді (2.9), де функції $p_k\left(x_k\right)$ невід'ємні й задовольняють умову

$$\int_{-\infty}^{\infty} p_k\left(x_k\right)\, dx_k = 1,$$

то випадкові величини ξ_k $(k = 1, 2, \ldots, n)$ незалежні і $p_k\left(x_k\right)$ є щільністю розподілу величини ξ_k.

Приклад 2.2. Якщо величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні й мають щільності розподілу

$$p_k\left(x_k\right) = \frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{(x_k-a_k)^2}{2\sigma_k^2}} \,(1 \leq k \leq n),$$

то n-вимірна щільність розподілу величини $\left(\xi_1, \xi_2, \ldots, \xi_n\right)$ дорівнює

$$p\left(x_1, x_2, \ldots, x_n\right) = \frac{(2\pi)^{-\frac{n}{2}}}{\sigma_1 \sigma_2 \cdots \sigma_n}\, e^{-\frac{1}{2}\sum\limits_{k=1}^{n} \frac{(x_k-a_k)^2}{2\sigma_k^2}}.$$

При $n = 2$ ця формула набуває такого вигляду:

$$p\left(x_1,\, x_2\right) = \frac{1}{2\pi\sigma_1\sigma_2}\, e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2} - \frac{(x_2-a_2)^2}{2\sigma_2^2}}.$$

Порівняння цієї функції зі щільністю двовимірного нормального закону показує, що для незалежних випадкових величин ξ_1 та ξ_2 параметр r дорівнює 0.

2.10 Функції від випадкових величин

Властивості функцій розподілу дають змогу приступити до розв'язування такої задачі: знаючи функцію розподілу $F(x_1, x_2, \dots, x_k)$ сукупності випадкових величин $\xi_1, \xi_2, \dots, \xi_k$, визначити функцію розподілу $\Phi(y_1, y_2, \dots, y_n)$ величин

$$\eta_1 = f_1(\xi_1, \xi_2, \dots, \xi_k), \eta_2 = f_2(\xi_1, \xi_2, \dots, \xi_k), \dots, \eta_n = f_n(\xi_1, \xi_2, \dots, \xi_k).$$

Загальне розв'язання цієї задачі дуже просте, але вимагає розширення поняття інтеграла. Далі обмежимося розглядом важливих для практики випадків: дискретних і неперервних випадкових величин.

Розглянемо спочатку випадок, коли сукупність величин $\xi_1, \xi_2, \dots, \xi_k$ має щільність розподілу ймовірностей $p(x_1, x_2, \dots, x_k)$.

Із попереднього очевидно, що шукана функція розподілу $\Phi(y_1, \dots, y_n)$ задається рівністю

$$\Phi(y_1, y_2, \dots, y_n) = \int \underset{D}{\dots} \int p(x_1, x_2, \dots, x_k)\, dx_1 dx_2 \dots dx_k,$$

де область інтегрування D визначається нерівностями

$$f_i(x_1, x_2, \dots, x_k) < y_i \quad (i = 1, 2, \dots, n).$$

Природно, що для випадку дискретних випадкових величин розв'язання дається через k -вимірну суму, також поширену на область D.

Тільки що зроблене загальне зауваження відносно розв'язання поставленої нами задачі застосуємо до кількох важливих окремих випадків.

2.11 Функція розподілу суми випадкових величин

Нехай щільність розподілу сукупності випадкових величин $\xi_1, \xi_2, ..., \xi_k$ є $p(x_1, x_2, \dots, x_k)$. Треба знайти функцію розподілу суми

$$\eta = \xi_1 + \xi_2 + ... + \xi_k.$$

Шукана функція розподілу дорівнює ймовірності потрапляння вектора $(\xi_1, \xi_2, ..., \xi_k)$ у півпростір

$$\{(x_1, x_2, \dots, x_k) : x_1 + x_2 + ... + x_k < x\}$$

і дорівнює інтегралу

$$\Phi\,(x) = \int \underset{\sum x_i < x}{\cdots} \int p\ (x_1,\, x_2, \ldots,\, x_k)\ dx_1\, dx_2 \ldots dx_k.$$

Розглянемо докладніше випадок $k = 2$. Попередня формула набуває в цьому випадку такого вигляду:

$$\Phi\,(x) = \iint\limits_{x_1 + x_2 < x} p\,(x_1,\, x_2)\ dx_1\, dx_2 = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{x - x_2} p\,(x_1,\, x_2)\ dx_1\, dx_2. \tag{2.10}$$

Якщо величини ξ_1 та ξ_2 незалежні, то

$$p\ (x_1,\, x_2) = p\ (x_1) \cdot p\ (x_2)$$

і рівність (2.10) переписується так:

$$\Phi\,(x) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{x - x_2} p_1\,(x_1)\ p_2\,(x_2)\ dx_1 dx_2 = \int\limits_{-\infty}^{\infty} F_1\,(x - x_2)\ p_2\,(x_2)\ dx_2. \tag{2.11}$$

Якщо інтегрування провести спочатку за змінними x_2, то формула набуває такого вигляду:

$$\Phi\,(x) = \int\limits_{-\infty}^{\infty} F_2\,(x - x_1)\ p_1\,(x_1)\ dx_1. \tag{2.12}$$

Диференціювання формул (2.11) або (2.12) приводить нас до рівностей

$$p\ (x) = \int\limits_{-\infty}^{\infty} p_1\,(x - z)\ p_2\,(z)\ dz = \int\limits_{-\infty}^{\infty} p_2\,(x - z)\ p_1\,(z)\ dz. \tag{2.13}$$

Так само диференціювання (2.10) приводить нас до рівностей

$$p\ (x) = \int\limits_{-\infty}^{\infty} p\ (x - z,\, z)\ \ dz = \int\limits_{-\infty}^{\infty} p\ (z,\, x - z)\ dz. \tag{2.14}$$

Останні рівності показують, що коли доданки мають щільність розподілу ймовірностей, то їхня сума також має щільність розподілу.

Розглянемо приклади.

2.11.1 Сума розподілів Пуассона

Нехай випадкові величини ξ_1 та ξ_2 незалежні й розподілені за законом Пуассона з параметрами λ_1 та λ_2 відповідно. Тобто $\xi_1 \simeq Pois(\lambda_1)$ та $\xi_2 \simeq Pois(\lambda_2)$.

Випадкова величина ξ розподілені за законом Пуассона з параметром λ приймає значення $0, 1, 2, \ldots, n, \ldots$ з імовірностями

$$p_k = P\{\xi = k\} = \frac{e^{-\lambda}\lambda^k}{k!},\ k = 0, 1, 2, \ldots$$

Знайдемо розподіл величини $\eta = \xi_1 + \xi_2$.

Формула для підрахунку розподілу суми двох нехалежних дискретних випадкових величин має вигляд

$$P\{\xi_1 + \xi_2 = n\} = \sum_{m=-\infty}^{\infty} P(\xi_1 = m)P(\xi_2 = n - m).$$

Враховуючи те, що випадкові величини розподілені за законом Пуассона, матимемо

$$P\{\xi_1 + \xi_2 = n\} = \sum_{m=0}^{n} \frac{e^{-\lambda_1}{\lambda_1}^m}{m!}\frac{e^{-\lambda_2}{\lambda_2}^{n-m}}{(n-m)!} =$$

$$= e^{-(\lambda_1+\lambda_2)}\frac{1}{n!}\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}{\lambda_1}^m{\lambda_2}^{n-m} =$$

$$= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}.$$

Можемо зробити висновок, що сума двох незалежних дискретних випадкових величин, розподілених за законом Пуассона з параметрами λ_1 та λ_2 відповідно, розподілена за законом Пуассона з параметром $\lambda_1 + \lambda_2$.

Отже маємо такий результат. Якщо $\xi_1 \simeq Pois(\lambda_1)$ та $\xi_2 \simeq Pois(\lambda_2)$ незалежні, то $\xi_1 + \xi_2 \simeq Pois(\lambda_1 + \lambda_2)$.

2.11.2 Сума рівномірних розподілів

Нехай ξ_1 та ξ_2 незалежні й рівномірно розподілені в інтервалі $[a,\, b]$ випадкові величини. Знайдемо розподіл величини $\eta = \xi_1+\xi_2$. Щільність розподілу ймовірностей ξ_1 та ξ_2 дорівнює

$$p_1(x) = p_2(x) = \begin{cases} 0, & \text{якщо } x < a \text{ або } \quad \mathrm{x > b}, \\ \frac{1}{b-a}, & \qquad \text{якщо } \mathrm{a \le x \le b}. \end{cases}$$

За формулою (15) знаходимо, що

$$p_{\eta}(x) = \int_{a}^{b} p_1(z) \cdot p_2(x-z)\, dz = \frac{1}{b-a} \int_{a}^{b} p_2(x-z)\, dz.$$

Із того, що при $x < 2a$ маємо

$$x - z < 2a - z < a,$$

а при $x > 2b$ маємо

$$x - z > 2b - z > b,$$

виводимо, що при $x < 2a$ і $x > 2b$ маємо

$$p_{\eta}(x) = 0.$$

Нехай тепер $2a < x < 2b$. Підінтегральна функція відмінна від нуля лише за тих значень z, які задовольняють нерівність

$$a < x - z < b,$$

або, що те саме, нерівність

$$x - b < z < x - a.$$

Через те, що при $x > 2a$, маємо $x - a > a$.. Очевидно, що $x - a \leq b$ при $x \leq a + b$. Отже, якщо $2a < x \leq a + b$, то

$$p_{\eta}(x) = \int_{a}^{x-a} \frac{dz}{(b-a)^2} = \frac{x-2a}{(b-a)^2}.$$

Так само при $a + b < x \leq 2b$ маємо

$$p_{\eta}(x) = \int_{x-b}^{b} \frac{dz}{(b-a)^2} = \frac{2b-x}{(b-a)^2}.$$

Зібравши разом отримані результати, знаходимо, що

$$p_{\eta}(x) = \begin{cases} 0, & \text{при } x \leq 2a \text{ або } x > 2b, \\ \frac{x-2a}{(b-a)^2}, & \text{при } 2a < x \leq a+b, \\ \frac{2b-x}{(b-a)^2}, & \text{при } a+b < x \leq 2b. \end{cases}$$

Функція $p_n(x)$ - це щільність розподілу Сімпсона.

Обчислення значно спрощується, якщо при розв'язанні цього прикладу користуватися не формулами, а геометричними міркуваннями. Зобразимо, як звичайно, ξ_1 та ξ_2 як прямокутні координати на площині. Тоді ймовірність нерівності $\xi_1 + \xi_2 < x$ при $2a < x \le a + b$ дорівнює ймовірності потрапляння у двічі заштрихований прямокутний трикутник (рис. 13). Вона, як легко підрахувати, дорівнює

$$F_\eta(x) = \frac{(x-2a)^2}{2(b-a)^2}.$$

При $a+b < x \le 2b$ ймовірність нерівності $\xi_1+\xi_2 < x$ дорівнює ймовірності потрапляння в усю заштриховану фігуру. Ця ймовірність дорівнює

$$F_\eta(x) = 1 - \frac{(2b-x)^2}{2(b-a)^2}.$$

Диференціювання приводить до щільності **розподілу Сімпсона.**

2.11.3 Сума нормальних розподілів

Двомірна випадкова величина (ξ_1, ξ_2) розподілена за нормальним законом

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left\{\frac{(x-a)^2}{\sigma_1^2} - 2r\frac{(x-a)(y-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{\sigma_2^2}\right\}}.$$

Знайти функцію розподілу суми $\eta = \xi_1 + \xi_2$.

Згідно з формулою (2.14)

$$p_\eta(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2(1-r^2)}\left\{\frac{(z-a)^2}{\sigma_1^2} - 2r\frac{(z-a)(x-z-b)}{\sigma_1\sigma_2} + \frac{(x-z-b)^2}{\sigma_2^2}\right\}} dz.$$

Позначимо для скорочення $v = x - a - b$, $u = z - a$. Тоді

$$p_\eta(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2(1-r^2)}\left\{\frac{u^2}{\sigma_1^2} - 2r\frac{u(v-u)}{\sigma_1\sigma_2} + \frac{(v-u)^2}{\sigma_2^2}\right\}} du.$$

Оскільки

$$\frac{u^2}{\sigma_1^2} - 2r\frac{u\,(\upsilon-u)}{\sigma_1\sigma_2} + \frac{(\upsilon-u)^2}{\sigma_2^2} = u^2\frac{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}{\sigma_1^2\sigma_2^2} - 2u\upsilon\frac{\sigma_1+r\sigma_2}{\sigma_1\sigma_2^2} + \frac{\upsilon^2}{\sigma_2^2} =$$
$$= \left[u\frac{\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}}{\sigma_1\sigma_2} - \frac{\upsilon}{\sigma_2}\frac{\sigma_1+r\sigma_2}{\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}}\right]^2 + \frac{\upsilon^2}{\sigma_2^2}\left[1 - \frac{(\sigma_1+r\sigma_2)^2}{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}\right] =$$
$$= \left[u\frac{\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}}{\sigma_1\sigma_2} - \frac{\upsilon(\sigma_1+r\sigma_2)}{\sigma_2\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}}\right]^2 + \frac{\upsilon^2\left(1-r^2\right)}{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}\,,$$

то, ввівши позначення

$$t = \frac{1}{\sqrt{1-r^2}}\left[u\frac{\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}}{\sigma_1\sigma_2} - \frac{\upsilon\,(\sigma_1+r\sigma_2)}{\sigma_2\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}}\right],$$

приведемо вираз для $p_\eta\,(x)$ до вигляду

$$p_\eta\,(x) = \frac{e^{-\frac{\upsilon^2}{2\left(\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2\right)}}}{2\pi\,\sqrt{\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2}} \cdot \int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt\,.$$

Через те, що $\upsilon = x - a - b$ та

$$\int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt = \sqrt{2\pi},$$

маємо

$$p_\eta\,(x) = \frac{1}{\sqrt{2\pi\,\left(\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2\right)}}e^{-\frac{(x-a-b)^2}{2\left(\sigma_1^2+\sigma_2^2+2r\sigma_1\sigma_2\right)}}.$$

Зокрема, якщо випадкові величини ξ_1 та ξ_2 незалежні, то $r = 0$ і формула для $p_\eta\,(x)$ набуває вигляду

$$p_\eta\,(x) = \frac{1}{\sqrt{2\pi\,\left(\sigma_1^2+\sigma_2^2\right)}}e^{-\frac{(x-a-b)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}}.$$

Ми маємо дуже важливий результат: сума нормально розподілених випадкових величин розподілена за нормальним законом.

Зауваження 2.7. Зауважимо, що у випадку, коли доданки незалежні, має місце й обернене твердження (теорема Г. Крамера): якщо сума двох незалежних випадкових величин розподілена за нормальним законом, то кожний доданок також розподілений за нормальним законом.

2.11.4 Сума гама розподілів

Нехай незалежні випадкові величини ξ_1 та ξ_2 мають гама розподіли з параметрами (α_1, β) та (α_2, β), тобто $\xi_1 \simeq \Gamma(\alpha_1, \beta)$ та $\xi_2 \simeq \Gamma(\alpha_2, \beta)$. Покажемо, що випадкова величина $\eta = \xi_1 + \xi_2$ має гама розподіл з параметрами $(\alpha_1 + \alpha_2, \beta)$, тобто $\xi_1 + \xi_2 \simeq \Gamma(\alpha_1 + \alpha_2, \beta)$.

Випадкова невід'ємна величина ξ має гама розподіл з параметрами (α, β) (позначення $\xi \simeq \Gamma(\alpha, \beta)$), якщо її щільність розподілу має вигляд

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0. \end{cases}$$

Щільність розподілу суми двох незалежних випадкових величин обчислюється за формулою

$$p_\eta(x) = \int_{-\infty}^{\infty} p_1(x-y) p_2(y) dy.$$

Для незалежних випадкових величин ξ_1 та ξ_2 які мають гама розподіли з параметрами (α_1, β) та (α_2, β) матимемо $x > 0$

$$p_\eta(x) = \int_0^x \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)} (x-y)^{\alpha_1 - 1} e^{-\beta(x-y)} \frac{\beta^{\alpha_2}}{\Gamma(\alpha_2)} y^{\alpha_2 - 1} e^{-\beta y} dy =$$

$$= \frac{\beta^{\alpha_1 + \alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta x} \int_0^x (x-y)^{\alpha_1 - 1} y^{\alpha_2 - 1} dy.$$

Заміна $y = tx$ приводить до такого виразу

$$p_\eta(x) = \frac{\beta^{\alpha_1 + \alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x^{\alpha_1 + \alpha_2 - 1} e^{-\beta x} \int_0^1 (1-t)^{\alpha_1 - 1} t^{\alpha_2 - 1} dt.$$

Враховуючи, що

$$\int_0^1 (1-t)^{\alpha_1 - 1} t^{\alpha_2 - 1} dt = B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)},$$

отримаємо при $x > 0$ наступний вираз

$$p_\eta(x) = \frac{\beta^{\alpha_1 + \alpha_2}}{\Gamma(\alpha_1 + \alpha_2)} x^{\alpha_1 + \alpha_2 - 1} e^{-\beta x},$$

а це є щільність гама розпроділу з параметрами $(\alpha_1 + \alpha_2, \beta)$.

Отже маємо такий результат. Якщо $\xi_1 \simeq \Gamma(\alpha_1, \beta)$ та $\xi_2 \simeq \Gamma(\alpha_2, \beta)$ незалежні, то $\xi_1 + \xi_2 \simeq \Gamma(\alpha_1 + \alpha_2, \beta)$.

2.11.5 Розподіл χ_n^2

Нехай ξ_1, $\xi_2, ..., \xi_n$ – незалежні випадкові величини, розподілені за одним і тим самим нормальним законом з параметрами a, σ^2. Розподіл випадкової величини

$$\chi_n^2 = \frac{1}{\sigma^2}\sum_{k=1}^{n}(\xi_k - a)^2$$

має назву χ_n^2– розподіл.

Щоб обчислити щільність випадкової величини χ_n^2 зауважимо, що випадкова величина $\eta = \xi^2$, де $\xi \sim N(0,1)$ - нормально розподілена випадкова величина з параметрами $a = 0, \sigma = 1$, має гама розподіл з параметрами $\alpha = 1/2, \beta = 1/2$. Дійсно

$$P\{\xi^2 < x\} = P\{|\xi| < \sqrt{x}\} = \frac{2}{\sqrt{2\pi}}\int_0^{\sqrt{x}} e^{-\frac{u^1}{2}}du.$$

Тому щільність випадкової величини $\eta = \xi^2$ дорівнює

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

Це щільність гама розподілу з параметрами $\alpha = 1/2, \beta = 1/2$. Тобто $\eta = \xi^2 \simeq \Gamma(1/2, 1/2)$.

Користуючись результатами попереднього прикладу, отримаємо що щільність випадкової величини χ_n^2 є гама розподіл з параметрами $\alpha = n/2, \beta = 1/2$. Тобто $\xi_n^2 \simeq \Gamma(n/2, 1/2)$.

Щільність розподілу випадкової величини χ_n^2 дорівнює

$$p\left(x\right) = \begin{cases} 0, & x < 0, \\ \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}}x^{\frac{n}{2}-1}e^{-\frac{1}{2}x}, & x \geq 0. \end{cases}$$

При $n = 3$ дістанемо відомий закон **розподілу Максвелла** випадкової величини

$$\chi_n^2 = \frac{\left(\xi_1 - a\right)^2}{\sigma^2} + \frac{\left(\xi_2 - a\right)^2}{\sigma^2} + \frac{\left(\xi_3 - a\right)^2}{\sigma^2}.$$

2.11.6 Розподіл χ_n

Розподіл випадкової величини

$$\chi_n = \sqrt{\frac{1}{\sigma^2}\sum_{k=1}^{n}(\xi_k - a)^2},$$

що має назву χ-розподіл, відіграє важливу роль у математичній статистиці.

Обчислимо функцію розподілу випадкової величини

$$\zeta = \chi/\sqrt{n}.$$

Очевидно, що для від'ємних значень аргументу функція розподілу $\Phi\ (y)$ величини ζ дорівнює нулю. Для додатних значень у величина $\Phi\ (y)$ дорівнює ймовірності потрапляння точки $(\xi_1, \xi_2, ..., \xi_k)$ всередину кулі

$$\sum_{k=1}^{n}(x_k - a)^2 = ny^2\sigma^2.$$

Таким чином,

$$\Phi\ (y) = \int \underset{\sum x_k^2 < ny^2}{\dots} \int \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum\limits_{i=1}^{n}\frac{x_i^2}{2}}\, dx_1\, dx_2 \dots dx_n.$$

Перейдемо до сферичних координат, тобто зробимо заміну

$$x_1 = \rho\, \cos\theta_1\, \cos\theta_2 \cdots \cos\theta_{n-1},$$

$$x_2 = \rho\, \cos\theta_1\, \cos\theta_2 \cdots \sin\theta_{n-1},$$

$$\dots\dots\dots\dots$$

$$x_n = \rho\, \sin\theta_1.$$

Приведемо $\Phi\ (y)$ до вигляду

$$\Phi\ (y) = C_n \int\limits_0^{y\sqrt{n}} e^{-\frac{\rho^2}{2}}\rho^{n-1}\, d\rho,$$

де

$$C_n = \int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \ldots \int\limits_{0}^{2\pi} \left(\frac{1}{\sqrt{2\pi}}\right)^n \cos^{n-2}\theta_1 \ldots \cos\theta_{n-1}\, d\theta_{n-1} \ldots d\theta_1$$

– стала, що залежить лише від n. Цю сталу легко обчислити, користуючись рівністю

$$\Phi\,(\infty) = 1 = C_n \int\limits_0^\infty e^{-\frac{\rho^2}{2}} \rho^{n-1}\, d\rho = \left(\sqrt{2}\right)^{n-2} C_n \int\limits_0^\infty e^{-z} z^{\frac{n-2}{2}}\, dz =$$

$$\left(\sqrt{2}\right)^{n-2} C_n\, \Gamma\left(\frac{n}{2}\right).$$

Звідси знаходимо, що

$$C_n \;= \frac{1}{\Gamma\left(\frac{n}{2}\right)\,\left(\sqrt{2}\right)^{n-2}}\,,$$

і, отже,

$$\Phi\,(y) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \int\limits_0^{\frac{ny^2}{2}} e^{-z} z^{\frac{n}{2}-1} dz.$$

Щільність розподілу ζ для $y \geq 0$ дорівнює

$$\varphi\,(y) = \frac{\sqrt{2n}}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{y\sqrt{n}}{\sqrt{2}}\right)^{n-1} e^{-\frac{y^2 n}{2}}. \qquad (2.15)$$

Звідси, зокрема при $n = 1$, ми дістанемо, природно, щільність розподілу, що дорівнює подвійній щільності вихідного нормального закону:

$$\varphi\,(y) = \sqrt{\frac{2}{\pi}}\, e^{-\frac{y^2}{2}} \qquad (y \geq 0)\,.$$

Із формули (2.15) неважко вивести щільність розподілу випадкової величини χ_n^2 з n степенями вільності ($y \geq 0$):

$$\varphi\,(y) = \frac{1}{2^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}},$$

2.12 Функція розподілу частки

Нехай щільність розподілу випадкового вектора (ξ, η) дорівнює $p(x, y)$ і нам потрібно знайти функцію розподілу величини $\zeta = \frac{\xi}{\eta}$.

Згідно з означенням

$$F_\zeta(x) = P\left\{\frac{\xi}{\eta} < x\right\}.$$

Нехай ξ та η розглядаються як координати точки на площині, тоді $F_\zeta(x)$ дорівнює ймовірності того, що точка (ξ, η) потрапить в область, координати точок якої задовольняють нерівність $\frac{\xi}{\eta} < x$.

Згідно із загальною формулою шукана ймовірність дорівнює

$$F_\zeta(x) = \int\limits_0^\infty \int\limits_{-\infty}^{zx} p(y,z)\, dy\, dz + \int\limits_{-\infty}^0 \int\limits_{zx}^\infty p(y,z)\, dy\, dz. \tag{2.16}$$

Звідси випливає, що коли ξ та η незалежні, а $p_1(x)$ та $p_2(x)$ – їхні щільності розподілу, то

$$F_\zeta(x) = \int\limits_0^\infty F_1(xz)\, p_2(z)\, dz + \int\limits_{-\infty}^0 (1 - F_1(xz))\, p_2(z)\, dz.$$

Продиференціювавши (2.16), знаходимо

$$p_\zeta(x) = \int\limits_0^\infty zp(xz, z)\, dz - \int\limits_{-\infty}^0 zp(xz, z)\, dz. \tag{2.17}$$

Зокрема, якщо величини ξ та η незалежні, то

$$p_\zeta(x) = \int\limits_0^\infty zp_1(zx)\, p_2(z)\, dz - \int\limits_{-\infty}^0 zp_1(zx)\, p_2(z)\, dz. \tag{2.18}$$

2.12.1 Розподіл частки двох нормально розподілених аеличин

Випадкова величина (ξ, η) розподілена за нормальним законом

$$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\frac{x^2}{\sigma_1^2} - 2r\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right]}.$$

Знайти функцію розподілу частки $\zeta = \frac{\xi}{\eta}$.

За формулою (2.17)

$$p_\zeta(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \left[\int_0^\infty z\, e^{-\frac{z^2}{2(1-r^2)} \cdot \frac{\sigma_2^2 x^2 - 2r\sigma_1\sigma_2 x + \sigma_1^2}{\sigma_1^2\sigma_2^2}} dz - \right.$$

$$\left. - \int_{-\infty}^0 z\, e^{-\frac{z^2}{2(1-r^2)} \cdot \frac{\sigma_2^2 x^2 - 2r\sigma_1\sigma_2 x + \sigma_1^2}{\sigma_1^2\sigma_2^2}} dz \right] =$$

$$= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_0^\infty z\, e^{-\frac{z^2}{2(1-r^2)} \cdot \frac{\sigma_2^2 x^2 - 2r\sigma_1\sigma_2 x + \sigma_1^2}{\sigma_1^2\sigma_2^2}} dz.$$

Замінимо змінну інтегрування, поклавши

$$u = \frac{z^2}{2(1-r^2)} \frac{\sigma_2^2 x^2 - 2r\sigma_1\sigma_2 x + \sigma_1^2}{\sigma_1^2\sigma_2^2}.$$

Вираз $p_\zeta(x)$ при цьому матиме такий вигляд:

$$p_\zeta(x) = \frac{\sigma_1\sigma_2\sqrt{1-r^2}}{\pi(\sigma_2^2 x^2 - 2r\sigma_1\sigma_2 x + \sigma_1^2)} \int_0^\infty e^{-u} du = \frac{\sigma_1\sigma_2\sqrt{1-r^2}}{\pi(\sigma_2^2 x^2 - 2r\sigma_1\sigma_2 x + \sigma_1^2)}.$$

Якщо, зокрема, величини ξ та η незалежні, то щільність набуває вигляду

$$p_\zeta(x) = \frac{\sigma_1\sigma_2}{\pi(\sigma_1^2 + \sigma_2^2 x^2)}.$$

Щільність розподілу величини ζ - це щільність **розподілу Коші.**

2.12.2 Розподіл Стьюдента

Знайти функцію розподілу частки $\zeta = \frac{\xi}{\eta}$, де ξ та η – незалежні величини, ξ розподілена за нормальним законом

$$p_\xi(x) = \frac{1}{\sigma}\sqrt{\frac{n}{2\pi}}\, e^{-\frac{nx^2}{2\sigma^2}},$$

а η – за законом $\chi/\sqrt{n}$:

$$p_\eta(x) = \frac{\sqrt{2n}}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{y\sqrt{n}}{\sqrt{2}}\right)^{n-1} e^{-\frac{ny^2}{2}}.$$

Згідно із формулою (2.18) маємо

$$p_\zeta(x) = \int\limits_0^\infty z\sqrt{\frac{n}{2\pi}}\, e^{-\frac{nz^2x^2}{2}} \cdot \frac{\sqrt{2n}}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{z\sqrt{n}}{\sqrt{2}}\right)^{n-1} e^{-\frac{nz^2}{2}}\, dz =$$

$$= \frac{1}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)} \int\limits_0^\infty \left(\frac{z\sqrt{n}}{\sqrt{2}}\right)^{n-1} e^{-\frac{nz^2}{2}\left(1+x^2\right)} nz\, dz.$$

Зробивши заміну

$$u = \frac{nz^2}{2}\left(1+x^2\right),$$

знаходимо, що

$$p_\zeta(x) = \frac{\left(x^2+1\right)^{-\frac{n+1}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)} \int\limits_0^\infty u^{\frac{n-1}{2}}\, e^{-u}\, du.$$

Звідси знаходимо щільність розподілу ймовірностей

$$p_\zeta(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)} \left(1+x^2\right)^{-\frac{n+1}{2}};$$

вона має назву **розподілу Стьюдента** (Стьюдент – псевдонім англійського математика-статистика У. Госсета, який уперше відкрив цей закон емпіричним шляхом).

При $n = 1$ розподіл Стьюдента – це розподіл Коші.

3 Послідовності незалежних дослідів

Розглянемо послідовність n незалежних дослідів, кожний із яких має k несумісних результатів $A_1^{(s)}$, $A_2^{(s)}, \ldots, A_k^{(s)}$ у досліді з номером $s = 1, 2, \ldots, n$, причому ймовірність результату $A_i^{(s)}$ не залежить від номера досліду s і дорівнює p_i $(i = 1, 2, \ldots, k\ )$, $\sum p_i = 1$. Ця схема була розглянута Я. Бернуллі при $k = 2$. У схемі Бернуллі $p_1 = p$, $p_2 = 1 - p = q$. Багато фактів, які виявлені при дослідженні схеми Бернуллі, пізніше були основою при вивченні більш складних схем. Пізніше переконаємося в цьому на прикладах закону великих чисел і теореми Муавра – Лапласа.

3.1 Біноміальний закон розподілу ймовірностей

Найпростіша задача, яка пов'язана зі схемою незалежних дослідів, полягає у визначенні ймовірності $P_n(m_1, m_2, \ldots, m_k)$ того, що при n дослідах відбудеться m_1 разів подія $A_1^{(s)}$, m_2 разів – подія $A_2^{(s)}$, $\ldots$, m_k разів – подія $A_k^{(s)}$. При цьому має виконуватися рівність

$$m_1 + m_2 + \cdots + m_k = n.$$

Обмежимося докладним дослідженням цієї задачі у випадку схеми Бернуллі. Для схеми Бернуллі позначимо подію $A_1^{(s)}$ через $A^{(s)}$, подію $A_2^{(s)}$ – через $\overline{A^{(s)}}$, число появ події $A^{(s)}$ – через μ та ймовірність $P_n(m_1, m_2)$ – через $P_n(m)$, де $m = m_1$. Ймовірність того, що події $A^{(s)}$ відбудуться при визначених m випробовуваннях (напр., при $s_1, s_2, \ldots, s_m$), а при інших $n - m$ дослідах не відбудуться, дорівнює $p^m q^{n-m}$. На основі теореми додавання шукана ймовірність $P_n(m)$ дорівнює сумі щойно обчислених імовірностей для всіх можливих способів, за яких подія $A^{(s)}$ відбудеться m разів і $n - m$ разів не відбудеться. Число таких способів дорівнює

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

отже, шукана ймовірність дорівнює

$$P_n\,(m) = C_n^m p^m q^{n-m}. \tag{3.1}$$

При цьому

$$\sum_{m=0}^{n} P_n\,(m) = 1.$$

Це співвідношення також можна вивести із рівності

$$\sum_{m=0}^{n} P_n\left(m\right) = \left(p+q\right)^n = 1^n = 1.$$

Зауважимо, що ймовірність $P_n(m)$ дорівнює коефіцієнту при x^m у розкладі бінома

$$\left(q+px\right)^n$$

за степенями x. З огляду на цю властивість сукупність ймовірностей $P_n(m)$ називають **біноміальним законом** розподілу ймовірностей.

Аналогічні міркування дозволяють переконатись, що

$$P_n\left(m_1, m_2, \ldots, m_k\right) = \frac{n!}{m_1!m_2!...m_k!} p_1^{m_1} p_2^{m_2} ... p_k^{m_k}, \qquad (3.2)$$

а також у тому, що ця ймовірність є коефіцієнтом при

$$x_1^{m_1} x_2^{m_2} \cdots x_k^{m_k}$$

у розкладі за степенями полінома

$$\left(p_1x_1 + p_2x_2 + ... + p_kx_k\right)^n.$$

Приклад 3.1. Імовірність того, що виріб певного виробника бракований, дорівнює 0,005. Яка ймовірність того, що з 10 000 наугад взятих виробів буде бракованих виробів: а) рівно 40, б) не більше 70.

У нашому прикладі $n = 10000$, $p = 0,005$, отже, за формулою (3.1) знаходимо, що а)

$$P_{10000}\left(40\right) = C_{10000}^{40} (0,995)^{9960} (0,005)^{40}.$$

Імовірність $P\{\mu \leq 70\}$ того, що кількість бракованих виробів буде не більше 70, дорівнює сумі ймовірностей того, що кількість бракованих виробів буде дорівнювати $0, 1, 2, \ldots, 70$. Отже,

$$P\left\{\mu \leq 70\right\} = \sum_{m=0}^{70} P_n\left(m\right) = \sum_{m=0}^{70} C_{10000}^{m} (0,995)^{10000-m} \cdot (0,005)^m.$$

Приклад показує, що безпосереднє обчислення ймовірностей за формулами (3.1) та (3.2) часто викликає великі технічні труднощі, тому виникає задача знайти прості наближені формули для ймовірностей $P_n(m)$, а також для сум вигляду

$$\sum_{m=s}^{t} P_n\left(m\right)$$

при великих значеннях n.

Дослідимо поведінку ймовірностей $P_n(m)$, $m = 1, 2, \ldots, n$ при сталих n.

Почнемо з вивчення $P_n(m)$ як функції m. Легко обчислити, що для $0 \leq m < n$

$$\frac{P_n\left(m+1\right)}{P_n\left(m\right)} = \frac{n-m}{m+1} \cdot \frac{p}{q},$$

звідки випливає, що

$$P_n\left(m+1\right) > P_n\left(m\right),$$

коли $(n-m)\,p > (m+1)\,q$, тобто коли $np - q > m$;

$$P_n\left(m+1\right) = P_n\left(m\right),$$

коли $m = np - q$ і, нарешті,

$$P_n\left(m+1\right) < P_n\left(m\right),$$

якщо $m > np - q$.

Ми бачимо, що ймовірність $P_n(m)$ зі збільшенням m спочатку зростає, потім досягає максимуму і при подальшому зростанні m зменшується. При цьому, якщо $np - q$ є цілим числом, то максимальне значення ймовірність $P_n(m)$ приймає для двох значень m, а саме для

$$m_0 = np - q$$

та

$$m'_0 = np - q + 1 = np + p.$$

Коли ж $np - q$ не є цілим числом, то ймовірність $P_n(m)$ досягає свого максимального значення при $m = [m_0 + 1]$, що дорівнює найбільшому цілому числу, меншому від $m_0 + 1$.

Число $m_0 + 1$ називають найімовірнішим значенням μ.

Ми бачимо, що коли $np - q$ є цілим числом, то μ має два найімовірніших значення: m_0 та $m'_0 = m_0 + 1$. Зауважимо, що коли $np - q < 0$, то

$$P_n(0) > P_n(1) > \cdots > P_n(n),$$

а якщо $np - q = 0$, то

$$P_n(0) = P_n(1) > \cdots > P_n(n).$$

Як побачимо далі, при великих значеннях n імовірності $P_n(m)$ для m, близьких до найімовірнішого значення μ, скільки-небудь помітно відрізняються від нуля. Цей факт пізніше доведемо у вигляді чітко сформульованої теореми. Проілюструємо сказане числовим прикладом. Приклад. Нехай $n = 50$, $p = \frac{1}{3}$. Найімовірніших значень маємо два: $m_0 = np - q = 16$ та $m_0 + 1 = 17$.

3.2 Локальна гранична теорема

При дослідженні числових прикладів попереднього параграфа ми зробили висновок, що при великих значеннях n m обчислення ймовірностей $P_n(m)$ за формулою (3.1) викликає значні труднощі. Необхідно встановити асимптотичні формули, які дали б змогу з досить великою точністю визначати ці ймовірності. Уперше формула такого роду була знайдена А. Муавром у 1730 р. для окремого випадку схеми Бернуллі при $p = q = \frac{1}{2}$, а пізніше узагальнена П. Лапласом на випадок довільного p, $0 < p < 1$. Ця формула була названа локальною теоремою Лапласа. Для встановлення історичної справедливості будемо її називати локальною теоремою Муавра – Лапласа.

Теорема 3.1. (Локальна теорема Муавра – Лапласа.) *Якщо ймовірність появи певної події A в n незалежних випробовуваннях стала і дорівнює p, $0 < p < 1$, то для ймовірності $P_n(m)$ того, що в цих випробовуваннях подія A здійсниться рівно m раз, задовольняє при $n \to \infty$ співвідношення*

$$\sqrt{npq}P_n(m) : \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \to 1, \tag{3.3}$$

яке має місце рівномірно для всіх m, для яких

$$x = \frac{m - np}{\sqrt{npq}} \tag{3.4}$$

лежить в якому-небудь скінченному інтервалі.

Доведення. Доведення спирається на відому з курсу аналізу формулу Стірлінга

$$s! = \sqrt{2\pi s} \cdot s^s e^{-s} e^{\theta_s},$$

в якій показник θ_s задовольняє нерівність

$$|\theta_s| \leq \frac{1}{12s}. \tag{3.5}$$

Зауважимо, що формулу (3.4) можна записати у вигляді

$$m = np + x\sqrt{npq}, \tag{3.6}$$

звідки

$$n - m = nq - x\sqrt{npq} \tag{3.7}$$

Останні 2 рівності дозволяють нам твердити, що коли x залишається обмеженим, то відповідні їм числа m та $n - m$ зростають до нескінченності при $n \to \infty$.

Застосовуючи формулу Стірлінга, дістанемо

$$P_n\left(m\right) = \frac{n!}{m!(n-m)!} p^m q^{n-m} =$$

$$= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{n}{m} p\right)^m \left(\frac{n}{n-m} q\right)^{n-m} e^{\theta},$$

де

$$\theta = \theta_n - \theta_m - \theta_{n-m} < \frac{1}{12}\left(\frac{1}{n} + \frac{1}{m} + \frac{1}{n-m}\right)$$

Якщо $a \leq x \leq b$, то величина θ рівномірно відносно x у цьому інтервалі прямує до нуля при $n \to \infty$, отже, множник e^{θ} за цих же умов рівномірно прямує до одиниці.

Розглянемо тепер величину

$$\ln A_n = \ln\left\{\left(\frac{n}{m} p\right)^m \left(\frac{n}{n-m} q\right)^{n-m}\right\} =$$

$$= -(np + x\sqrt{npq}) \ln\left(1 + x\sqrt{\frac{q}{np}}\right) -$$

$$-(nq - x\sqrt{npq})\ln\left(1 - x\sqrt{\frac{p}{nq}}\right).$$

Згідно з умовами теореми величини

$$\sqrt{\frac{q}{np}}, \quad \sqrt{\frac{p}{nq}},$$

при досить великих n можна зробити як завгодно малими, тому можемо скористатися розкладом у ряд Маклорена функцій

$$\ln\left(1 + x\sqrt{\frac{q}{np}}\right), \quad \ln\left(1 - x\sqrt{\frac{p}{nq}}\right).$$

Обмежившись двома першими доданками, знаходимо, що

$$\ln\left(1 + x\sqrt{\frac{q}{np}}\right) = x\sqrt{\frac{q}{np}} - \frac{1}{2}\frac{qx^2}{np} + O\left(\frac{1}{n^{3/2}}\right),$$

$$\ln\left(1 - x\sqrt{\frac{p}{nq}}\right) = -x\sqrt{\frac{p}{nq}} - \frac{1}{2}\frac{px^2}{nq} + O\left(\frac{1}{n^{3/2}}\right).$$

Оцінка останнього доданка рівномірна в довільному скінченному інтервалі змінної x. Таким чином

$$\ln A_n = -(np + x\sqrt{npq})\left(x\sqrt{\frac{q}{np}} - \frac{1}{2}\frac{qx^2}{np} + O\left(\frac{1}{n^{3/2}}\right)\right) -$$

$$-(nq - x\sqrt{npq})\left(-x\sqrt{\frac{p}{nq}} - \frac{1}{2}\frac{px^2}{nq} + O\left(\frac{1}{n^{3/2}}\right)\right) =$$

$$= -\frac{1}{2}x^2 + O\left(\frac{1}{\sqrt{n}}\right).$$

і, отже, рівномірно відносно x у скінченних інтервалах $a \le x \le b$ при $n \to \infty$ має місце співвідношення

$$A_n \to e^{-\frac{1}{2}x^2}.$$

Далі маємо

$$\sqrt{npq} \cdot \sqrt{\frac{n}{m(n-m)}} \to 1$$

при $n \to \infty$, причому рівномірно в кожному скінченному інтервалі змінної x . Отримані співвідношення доводять теорему. $\square$

Сформулюємо аналогічну теорему для загального випадку схеми незалежних випробовувань, Перед тим, як формулювати теорему, введемо деякі позначення. Покладемо $q_i = 1 - p_i$,

$$x_i = \frac{m_i - np_i}{\sqrt{np_i q_i}}, \quad i = 1, 2, \ldots, k. \tag{3.8}$$

Величина x_i залежить не тільки від i (тобто від p_i), але також від n та від m_i, проте для скорочення запису не вводимо нових індексів.

Теорема 3.2. (Локальна гранична теорема.) *Якщо ймовірності $p_1, ..., p_k$ настання відповідно подій $A_1^{(s)}, ..., A_k^{(s)}$ в s-му випробовуванні не залежать від номера випробовування та відмінні від 0 і 1 $(0 < p_i < 1, i = 1, 2, ..., k)$, то ймовірність $P_n\left(m_1, m_2, \ldots, m_k\right)$ того, що при n незалежних випробуваннях події $A_i^{(s)}$ $(i = 1, 2, ..., k)$ відбудуться m_i разів $(m_1 + m_2 + ... + m_k = n)$, задовольняє співвідношення*

$$\sqrt{n^{k-1}} P_n\left(m_1, m_2, ..., m_k\right) : \frac{e^{-\frac{1}{2}\sum\limits_{i=1}^{k} q_i x_i^2}}{(2\pi)^{\frac{k-1}{2}} \sqrt{p_1 p_2 ... p_k}} \to 1, \quad n \to \infty \tag{3.9}$$

рівномірно для всіх m_i $(i = 1, 2, ..., k$), для яких x_i лежать у скінченних інтервалах $a_i \le x_i \le b_i$.

Доведення цієї теореми аналогічне доведенню теореми Муавра - Лапласа.

При $k = 2$ отримаємо твердження теореми Муавра - Лапласа.

Тепер можемо довести до кінця обчислення в прикладі, розглянутому у попередньому параграфі.

Приклад 3.2. У прикладі попереднього параграфа ми мали визначити $P_n\left(m\right)$ при $n = 10000$, $m = 40$, $p = 0,005$. Згідно зі доведеною теоремою, маємо

$$P_n\left(m\right) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{1}{2}\left(\frac{m-np}{\sqrt{npq}}\right)^2}.$$

Для нашого прикладу

$$\sqrt{npq} = \sqrt{10000 \cdot 0,005 \cdot 0,995} = \sqrt{49,75} \approx 7,05$$

$$\frac{m - np}{\sqrt{npq}} \approx -1,42.$$

Отже,

$$P_n\left(m\right) \approx \frac{1}{7,05\sqrt{2\pi}} e^{-\frac{(1,42)^2}{2}}.$$

Функція

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

табульована. За таблицею значень цієї функції знаходимо, що

$$P_n\left(m\right) \approx \frac{0,1456}{7,05} = 0,0206.$$

Точні обчислення, без використовування теореми Муавра – Лапласа, дають

$$P_n\left(m\right) \approx 0,0197.$$

3.3 Інтегральна гранична теорема

Доведену локальну граничну теорему використаємо для доведення інтегральної граничної теореми. Почнемо з найпростішого окремого випадку цієї теореми – інтегральної теореми Муавра – Лапласа.

Теорема 3.3. (Інтегральна теорема Муавра – Лапласа.) *Якщо μ – це кількість появ події A при n незалежних випробовуваннях, у кожному з яких імовірність цієї події дорівнює p, $0 < \ < 1$, то рівномірно відносно a та b $(-\infty \leq a < b \leq +\infty)$ при $n \to \infty$ має місце співвідношення*

$$P\left\{a \leq \frac{\mu - np}{\sqrt{npq}} < b\right\} \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz. \qquad (3.10)$$

Доведення. Введемо для скорочення запису позначення

$$P_n\left(a, b\right) = P\left\{a \leq \frac{\mu - np}{\sqrt{npq}} < b\right\}.$$

Ця ймовірність, очевидно, дорівнює сумі $\sum P_n\left(m\right)$, яка охоплює всі ті значення m, для яких $a \leq x_m < b$, де

$$x_m = \frac{m - np}{\sqrt{npq}}.$$

Визначимо тепер функцію $y = \Pi_n(x)$ за формулою:

$$y = \begin{cases} 0, & x < x_0 = -\frac{np}{\sqrt{npq}}; \\ 0, & x \geq x_n + \frac{1}{\sqrt{npq}} = \frac{1+np}{\sqrt{npq}}; \\ \sqrt{npq}P_n\left(m\right), & x_m \leq x < x_{m+1} \; (m = 0, 1, ..., n)\,. \end{cases}$$

Очевидно, що ймовірність $P_n\left(m\right)$ дорівнює площі, обмеженій кривою $y = \Pi_n(x)$, віссю OX і ординатами в точках $x = x_m$ та $x = x_{m+1}$, тобто

$$P_n\left(m\right) = \sqrt{npq}P_n\left(m\right)\left(x_{m+1} - x_m\right) = \int\limits_{x_m}^{x_{m+1}} \Pi_n(x)\,dx.$$

Звідси випливає, що шукана ймовірність $P_n\left(a, b\right)$ дорівнює площі, що лежить між кривою $y = \Pi_n(x)$, віссю OX і ординатами в точках $x_{\underline{m}}$ та $x_{\overline{m}}$, де $x_{\underline{m}}$ та $x_{\overline{m}}$ визначаються нерівностями

$$a \leq x_{\underline{m}} < a + \frac{1}{\sqrt{npq}},$$

$$b \leq x_{\overline{m}} < b + \frac{1}{\sqrt{npq}}.$$

Отже,

$$P_n\left(a, b\right) = \int\limits_{x_{\underline{m}}}^{x_{\overline{m}}} \Pi_n(x)\,dx =$$

$$= \int\limits_a^b \Pi_n(x)\,dx + \int\limits_b^{x_{\overline{m}}} \Pi_n(x)\,dx - \int\limits_a^{x_{\underline{m}}} \Pi_n(x)\,dx.$$

Оскільки максимальне значення ймовірності $P_n\left(m\right)$ досягається коли $m_0 = \left[\left(n+1\right)p\right]$, то максимальне значення $\prod_n\left(x\right)$ досягається на інтервалі

$$\frac{m_0 - np}{\sqrt{npq}} \leq x \leq \frac{m_0 + 1 - np}{\sqrt{npq}} \leq \frac{2}{\sqrt{npq}}.$$

У цьому інтервалі має місце локальна теорема Муавра – Лапласа. Можна зробити висновок, що при всіх досить великих значеннях n

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

Звідси насамперед виводимо, що

$$|\rho_n| = \left|\int\limits_b^{x_{\overline{m}}} \Pi_n(x)\,dx - \int\limits_a^{x_{\underline{m}}} \Pi_n(x)\,dx\right| < \int\limits_b^{x_{\overline{m}}} \max \Pi_n(x)dx +$$
$$+ \int\limits_a^{x_{\underline{m}}} \max \Pi_n(x)dx < \sqrt{\tfrac{2}{\pi}}\left(x_{\overline{m}} - b + x_{\underline{m}} - a\right) \le 2\sqrt{\tfrac{2}{\pi npq}}$$

і, отже,

$$\lim_{n\to\infty} \rho_n = 0.$$

Таким чином, $P_n\,(a,b)$ відрізняється від $\int\limits_a^b \Pi_n(x)\,dx$ лише на нескінченно малу величину. Припустимо спочатку, що a і b скінченні числа. За цього припущення згідно з локальною теоремою при $a \le x_m < b$

$$\Pi_n(x_m) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_m^2}{2}} \left(1 + a_n\,(x_m)\right),$$

де $a_n\,(x_m) \to 0$ при $n \to \infty$ рівномірно відносно x_m. Очевидно, що при проміжних значеннях аргументу

$$\Pi_n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(1 + a_n\,(x)\right),$$

причому $\lim\limits_{n\to\infty} \max\limits_{a\le x<b} a_n\,(x) = 0$. Справді, при довільному m в інтервалі $x_m \le x < x_{m+1}$ маємо

$$\Pi_n(x) = \Pi_n(x_m) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(1 + a_n\,(x)\right),$$

де

$$a_n\,(x) = a_n\,(x_m) + \left(e^{\frac{x^2 - x_m^2}{2}} - 1\right)\left(a_n\,(x_m) + 1\right) - 1.$$

Оскільки

$$\frac{x^2 - x_m^2}{2} \le |x| \cdot |x - x_m| < \frac{\max\left(|a|\,, |b|\right)}{\sqrt{npq}},$$

то зрозуміло, що

$$\lim_{n\to\infty} \max_{a\le x<b} a_n\,(x) = 0.$$

Із сукупності доведених співвідношень отримуємо, що

$$P_n\,(a,b) = \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx + R_n,$$

де

$$R_n = \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} a_n(x)\, dx + \rho_n.$$

Оскільки

$$|R_n| \le \max |a_n(x)| \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx + \rho_n,$$

то з доведеного зрозуміло, що

$$\lim_{n\to\infty} R_n = 0.$$

При зроблених нами додаткових припущеннях теорему доведено.

Залишається відкинути ці обмеження. Для цього насамперед зауважимо, що

$$\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = 1.$$

Тому для довільного $\varepsilon > 0$можна вибрати настільки велике A, що

$$\frac{1}{\sqrt{2\pi}} \int\limits_{-A}^{A} e^{-\frac{z^2}{2}} dz > 1 - \frac{\varepsilon}{4},$$

$$\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{A} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int\limits_{A}^{\infty} e^{-\frac{z^2}{2}} dz < \frac{\varepsilon}{8}.$$

Виберемо далі згідно з доведеним настільки велике n, що при $-A \le a \le b < A$ матимемо

$$\left| P_n(a,b) - \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{z^2}{2}} dz \right| < \frac{\varepsilon}{4}.$$

Тоді очевидно, що

$$P_n(-A, A) > 1 - \frac{\varepsilon}{2},$$

$$P_n(-\infty, -A) + P_n(A, \infty) = 1 - P_n(-A, A) < \frac{\varepsilon}{2}.$$

Тепер легко буде довести, що при довільних a та b $(-\infty \leq a \leq b < \infty)$ має місце нерівність

$$\left| P_n(a,b) - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz \right| < \varepsilon,$$

що, очевидно, завершить доведення теореми Лапласа. Для цього необхідно розглянути окремі випадки розміщення на прямій точок a та b відносно $(-A, A)$. Розглянемо найскладніший випадок $(a \leq -A, b \geq A)$ (інші залишимо читачеві). У цьому випадку

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \left(\int_a^{-A} e^{-\frac{z^2}{2}} dz + \int_{-A}^{A} e^{-\frac{z^2}{2}} dz + \int_A^b e^{-\frac{z^2}{2}} dz \right),$$

$$P_n(a,b) = P_n(a,-A) + P_n(-A,A) + P_n(A,b),$$

а тому

$$\left| P_n(a,b) - \tfrac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz \right| \leq \left| P_n(a,-A) - \tfrac{1}{\sqrt{2\pi}} \int_a^{-A} e^{-\frac{z^2}{2}} dz \right| +$$

$$+ \left| P_n(-A,A) - \tfrac{1}{\sqrt{2\pi}} \int_{-A}^{A} e^{-\frac{z^2}{2}} dz \right| + \left| P_n(A,b) - \tfrac{1}{\sqrt{2\pi}} \int_A^b e^{-\frac{z^2}{2}} dz \right| \leq$$

$$\leq P_n(-\infty,-A) + \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{-A} e^{-\frac{z^2}{2}} dz + \left| P_n(-A,A) - \tfrac{1}{\sqrt{2\pi}} \int_{-A}^{A} e^{-\frac{z^2}{2}} dz \right| +$$

$$+ P_n(A,\infty) + \tfrac{1}{\sqrt{2\pi}} \int_A^{\infty} e^{-\frac{z^2}{2}} dz < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{4} + \tfrac{\varepsilon}{8} + \tfrac{\varepsilon}{8} = \varepsilon.$$

$\square$

Справджується така теорема.

Теорема 3.4. *У схемі послідовності незалежних випробувань, у кожному з яких можливі k результатів, причому ймовірність кожного з результатів не залежить від номера випробування і відмінна від 0 і від 1, для довільної області G гіперплощини, для якої $(k-1)$-вимірний об'єм її контуру дорівнює нулю, має місце співвідношення: при $n \to \infty$*

$$P_n(G) \to \sqrt{\frac{q_1 q_2 \ldots q_k}{(2\pi)^{k-1} \sum\limits_{i=1}^{k} p_i q_i}} \int_G e^{-\frac{1}{2} \sum\limits_{i=1}^{k} q_i x_i^2} dv, \qquad (3.11)$$

де dv означає елемент об'єму області G, і інтеграл береться по всій області G.

Доведення як за змістом, так і за своїм методом є майже копією міркувань, які було проведено при доведенні інтегральної теореми Муавра – Лапласа.

3.4 Застосування інтегральної теореми Муавра – Лапласа

Як перше застосування цієї теореми оцінимо ймовірність виконання нерівності

$$\left|\frac{\mu}{n} - p\right| < \varepsilon,$$

де $\varepsilon > 0$ – стала.

Маємо

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = P\left\{-\varepsilon\sqrt{\frac{n}{pq}} < \frac{\mu - np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}\right\}$$

і, отже, унаслідок інтегральної теореми Муавра – Лапласа

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz = 1.$$

Таким чином, для будь-якої сталої $\varepsilon > 0$ ймовірність виконання нерівності

$$\left|\frac{\mu}{n} - p\right| < \varepsilon$$

прямує до одиниці. Це показує, що частота настання події має тенденцію наближатися до ймовірності зі збільшенням числа випробувань.

Цей факт був уперше встановлений Я. Бернуллі. Він називається законом великих чисел, або теоремою Бернуллі.

Розглянемо тепер типові задачі, які приводять до теореми Муавра-Лапласа.

Проводиться n незалежних випробовувань, при кожному з яких імовірність настання події A дорівнює p.

I. Ставиться питання, чому дорівнює ймовірність того, що частота настання події A відхиляється від імовірності p не більше ніж на a. Ця ймовірність дорівнює

$$P\left\{\left|\frac{\mu}{n}-p\right|\leq a\right\}=P\left\{-a\cdot\sqrt{\frac{n}{pq}}\leq\frac{\mu-np}{\sqrt{npq}}\leq a\cdot\sqrt{\frac{n}{pq}}\right\}\approx$$
$$\approx\frac{1}{\sqrt{2\pi}}\int\limits_{-a\cdot\sqrt{\frac{n}{pq}}}^{a\cdot\sqrt{\frac{n}{pq}}}e^{-\frac{x^2}{2}}dx=\frac{2}{\sqrt{2\pi}}\int\limits_{0}^{a\cdot\sqrt{\frac{n}{pq}}}e^{-\frac{x^2}{2}}dx$$

II. Яке найменше число випробувань треба зробити, щоб імовірність того, що частота події A відхиляється від імовірності p не більше ніж на a, була не менша від β. Треба визначити n із нерівності

$$P\left\{\left|\frac{\mu}{n}-p\right|\leq a\right\}\geq\beta.$$

Імовірність, що стоїть у лівій частині нерівності, заміняємо наближено інтегралом за теоремою Муавра – Лапласа. Для визначення n отримуємо нерівність

$$\frac{2}{\sqrt{2\pi}}\int\limits_{0}^{a\cdot\sqrt{\frac{n}{pq}}}e^{-\frac{x^2}{2}}dx\geq\beta.$$

III. За даної імовірності β і кількості випробувань n треба визначити границю можливих змін $\left|\frac{\mu}{n}-p\right|$. Інакше кажучи, знаючи β і n, треба знайти a, для якого

$$P\left\{\left|\frac{\mu}{n}-p\right|<a\right\}=\beta.$$

Застосування інтегральної теореми Лапласа дає нам для визначення a рівність

$$\frac{2}{\sqrt{2\pi}}\int\limits_{0}^{a\cdot\sqrt{\frac{n}{pq}}}e^{-\frac{x^2}{2}}dx=\beta.$$

Чисельне розв'язання всіх розглянутих нами задач вимагає вміння обчислювати значення інтеграла

$$\Phi\left(x\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{x}e^{-\frac{z^2}{2}}dz \tag{3.12}$$

за довільних значень x і розв'язувати обернену задачу: за величиною інтеграла $\Phi\left(x\right)$ обчислювати відповідне значення аргументу x.

Для цих обчислень потрібні спеціальні таблиці, оскільки інтеграл (3.12) при $0 < x < \infty$ у скінченному вигляді в елементарних функціях не обчислюється.

Тепер можемо довести розв'язання розглянутих прикладів до числового результату.

Приклад 3.3. У прикладі 3.1 треба було знайти ймовірність того, що число бракованих виробів буде не більше 70, якщо ймовірність для кожного виробу бути бракованим дорівнює р = 0,005 і число виробів дорівнює 10 000. За щойно доведеною теоремою ця ймовірність дорівнює

$$
\begin{gathered}
P\{\mu \le 70\} = P\left\{-\tfrac{50}{\sqrt{49{,}75}} \le \tfrac{\mu - np}{\sqrt{npq}} \le \tfrac{20}{\sqrt{49{,}75}}\right\} = \\
= P\left\{-7,09 \le \tfrac{\mu - np}{\sqrt{npq}} \le 2,84\right\} \approx \\
\approx \tfrac{1}{\sqrt{2\pi}} \int\limits_{-7{,}09}^{2{,}84} e^{-\frac{z^2}{2}} dz = \Phi(2,84) - \Phi(-7,09) = \\
= \Phi(2,84) + \Phi(7,09) = 0,9975
\end{gathered}
$$

Значення функції $\Phi(x)$ при $x = 7,09$ у таблицях немає, ми замінили його 0,5, зробивши при цьому похибку меншу від $10^{\breve{}10}$.

3.5 Теорема Пуассона

При доведенні локальної теореми Муавра – Лапласа ми бачили, що асимптотичне вираження ймовірності $P_n(m)$ за допомогою функції

$$
\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}
$$

буде тим гірше, чим більше ймовірність p відрізняється від 0,5, тобто чим менші значення p або qдоводиться розглядати. Цю формулу також не можна використовувати при $= 0, q = 1$, а також при $q = 0, = 1$. Проте значна частина задач пов'язана з необхідністю обчислювати ймовірності $P_n(m)$ саме при малих значеннях p. Щоб у цьому випадку теорема Муавра – Лапласа дала результат з незначною похибкою, необхідно щоб кількість випробувань n була дуже великою. Таким чином, виникає потреба знайти асимптотичну формулу спеціально пристосовану для цього випадку. Така формула знайдена Пуассоном.

Теорема 3.5. (Теорема Пуассона.) *Якщо ймовірність p_n появи події A в n незалежних випробовуваннях мала, то ймовірність*

$P_n(\mu_n = m)$ того, що при n незалежних випробовуваннях подія A відбудеться m разів, асимптотично дорівнює

$$P_n(\mu_n = m) \sim \frac{a^m e^{-a}}{m!}, \quad m = 0, 1, \ldots \tag{3.13}$$

за умоваи $n \to \infty$, $p_n \to 0$ так що $np_n \to a$.

Доведення. Справді, ймовірність m-кратної появи події A при n незалежних випробовуваннях дорівнює

$$P_n(\mu_n = m) = C_n^m {p_n}^m (1 - p_n)^{n-m} =$$

$$= a_{nm} \frac{(np_n)^m}{m!} \left(1 - \frac{np_n}{n}\right)^n,$$

де

$$a_{nm} = (1 - p_n)^{-m} \prod_{1 \le i \le m} \left(1 - \frac{i}{n}\right) \to 1,\ n \to \infty.$$

Крім того,

$$\lim_{n\to\infty} (np_n)^m = a^m, \quad \lim_{n\to\infty} \left(1 - \frac{np_n}{n}\right)^n = e^{-a}$$

при сталому m.

Ці співвідношення завершують доведення теореми. □

Зауважимо, що теорема Пуассона має місце також у випадку, коли ймовірність події A в кожному випробуванні дорівнює нулю, у цьому випадку $a = 0$. Позначимо

$$P(m) = \frac{a^m}{m!} e^{-a}.$$

Розподіл імовірностей $P(m)$ називається законом Пуассона. Вивчимо поведінку $P(m)$ як функції m. Для цього розглянемо відношення

$$\frac{P(m)}{P(m-1)} = \frac{a}{m}.$$

Ми бачимо, що коли $m > a$, то

$$P(m) < P(m-1)$$

, якщо ж $m < a$, то

$$P(m) > P(m-1),$$

якщо, нарешті, $m = a$, то

$$P(m) = P(m-1).$$

Звідси виводимо, що величина $P(m)$ зростає при збільшенні m від 0 до $m_0 = [a]$; при подальшому зростанні m імовірність $P(m)$ зменшується. Якщо a – ціле число, то $P(m)$ має два максимальні значення: при $m_0 = a$ і при $m'_0 = a - 1$.

Приклад 3.4. Імовірність влучання в ціль при кожному пострілі дорівнює 0,001. Знайти ймовірність влучання в ціль двома і більше кулями, якщо число пострілів дорівнює 5 000.

Вважаючи кожний постріл за випробування, а влучання в ціль – за подію, можемо для обчислення ймовірності $P_n\{\mu \geq 2\}$ скористатися теоремою Пуассона. Згідно з указаною нами загальною схемою обчислень вважатимемо, що

$$p = 0,001 = \frac{a}{n} = \frac{a}{5000},$$

звідки $a = 5$. Шукана ймовірність дорівнює

$$P_n\{\mu \geq 2\} = \sum_{m=2}^{5000} P_n(m) = 1 - P_n(0) - P_n(1).$$

За теоремою Пуассона

$$P_n(0) \approx e^{-5}, P_n(1) \approx 5e^{-5}.$$

Отже,

$$P_n\{\mu \geq 2\} \approx 1 - 6e^{-5} \approx 0,9596.$$

Максимальне значення ймовірність $P_n(m)$ приймає при $m = 4$ і $m = 5$. Ці ймовірності рівні із точністю до четвертого десяткового знаку

$$P(4) = P(5) \approx 0{,}1751.$$

Обчислення за точною формулою дають із точністю до четвертого знака

$$P_{5000}(0) = 0{,}0071, \quad P_{5000}(1) = 0{,}0354\ ,$$

і, отже,

$$P_n\{\mu \geq 2\} = 0,9575.$$

Похибка від використання асимптотичної формули менша від $0,25\%$ обчислюваної величини.

4 Числові характеристики випадкових величин

У теорії ймовірностей та її застосуваннях велику роль відіграють деякі сталі числа, які дістаємо за певними правилами із законів розподілу випадкових величин. Серед цих сталих, які служать для числової характеристики випадкових величин, особливо важливі математичне сподівання, дисперсія та моменти різних порядків.

4.1 Математичне сподівання випадкових величин

Означення 4.1. Математичним сподіванням випадкової величини ξ називається інтеграл

$$\mathsf{E}\xi = \int\limits_{-\infty}^{\infty} x dF_{\xi}(x), \tag{4.1}$$

який домовимося позначати символом $\mathsf{E}\xi$.

Останнім часом у теорії ймовірностей, а також в її застосуваннях до техніки, фізики та біології починає поширюватись інша термінологія, і часто інтеграл (4.1) називають середнім значенням величини ξ. Далі будемо користуватися терміном "математичне сподівання", щоб уникнути можливої заміни цього поняття поняттям середнього, що використовується в статистиці. Якщо випадкова величина ξ неперервна і $p(x)$ її щільність розподілу, то

$$\mathsf{E}\xi = \int\limits_{-\infty}^{\infty} x\, p(x)\ dx \tag{4.2}$$

за умови що існує ітреграл

$$\mathsf{E}\xi = \int\limits_{-\infty}^{\infty} |x|\, p(x)\ dx$$

Якщо ж випадкова величина ξ дискретна, $x_1, x_2, \dots, x_n, \dots$ її можливі значення ξ та $p_1, p_2, \dots, p_n, \dots$ – відповідні їм ймовірності, то формула (4.1) набуває вигляду

$$\mathsf{E}\xi = \sum_{n=1}^{\infty} x_n p_n. \tag{4.3}$$

Отже, для дискретної випадкової величини математичне сподівання дорівнює сумі добутків значень цієї величини на їхні ймовірності. Наприклад, якщо можливі значення випадкової величини ξ будуть $x_1, x_2, \ldots, x_n$ та ймовірності цих значень усі рівні між собою $p_1 = p_2 = \ldots = p_n = \frac{1}{n}$, то математичне сподівання ξ є середнє арифметичне її можливих значень

$$\mathsf{E}\xi = \frac{1}{n}\sum_{k=1}^{n} x_k.$$

Зрозуміло, що математичне сподівання існує не для всіх випадкових величин: для його існування необхідна збіжність інтеграла (4.1) і до того ж абсолютна. Коли б інтеграл (4.1) збігався неабсолютно, то його значення залежало б від способу його обчислення, а тому математичне сподівання було б не стільки характеристикою випадкової величини, скільки б виразом волі обчислювача. Серед розглянутих прикладів випадкова величина, розподілена за законом Коші, не має математичного сподівання.

Розглянемо приклади обчислення математичних сподівань.

Приклад 4.1. Знайдемо математичне сподівання випадкової величини ξ яка має біноміальний розподід

$$P\,\{\,\xi = k\} = C_n^k p^k q^{n-k}, \quad k = 0, 1, \ldots, n.$$

Скористаємося формулою

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

Маємо

$$\mathsf{E}\xi = \sum_{k=0}^{n} k C_n^k p^k q^{n-k} = np\sum_{k=1}^{n} \frac{k}{n} C_n^k p^{k-1} q^{n-k} =$$

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

Отже

$$\mathsf{E}\xi = np.$$

Приклад 4.2. Знайдемо математичне сподівання випадкової величини ξ, розподіленої за геометричним законом:

$$P\{\xi = k\} = q^k p \quad k = 0, 1, 2, \ldots$$

Маємо

$$\mathsf{E}\xi = \sum_{k=0}^{\infty} k q^k p = qp \sum_{k=1}^{\infty} k q^{k-1} = \tfrac{q}{p}.$$

Отже

$$\mathsf{E}\xi = \frac{q}{p}.$$

Приклад 4.3. Знайдемо математичне сподівання випадкової величини ξ, розподіленої за законом Пуассона:

$$P\{\xi = k\} = \frac{a^k e^{-a}}{k!} \quad (k = 0, 1, 2, \ldots).$$

Маємо

$$\mathsf{E}\xi = \sum_{k=0}^{\infty} k \frac{a^k e^{-a}}{k!} = \sum_{k=1}^{\infty} k \frac{a^k e^{-a}}{k!} = a e^{-a} \sum_{k=1}^{\infty} \frac{a^{k-1}}{(k-1)!} =$$

$$= a e^{-a} \cdot e^a = a.$$

Отже

$$\mathsf{E}\xi = a.$$

Приклад 4.4. Знайдемо математичне сподівання випадкової величини ξ, розподіленої за нормальним законом:

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

За формулою (2) знаходимо, що

$$\mathsf{E}\xi = \int_{-\infty}^{\infty} x \, \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{(x-a)^2}{2\sigma^2}} \, dx.$$

Заміною $z = \frac{x-a}{\sigma}$ приводимо цей інтеграл до вигляду

$$\mathsf{E}\xi = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma x + a) \; e^{-\frac{x^2}{2}} \, dx = \tfrac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \, e^{-\frac{x^2}{2}} \, dx +$$

$$+ \tfrac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \, dx.$$

Оскільки
$$\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}, \ \int\limits_{-\infty}^{\infty} x\,e^{-\frac{x^2}{2}}\,dx = 0,$$
то $\mathsf{E}\xi = a$.

Дістаємо важливий результат, який з'ясовує сенс одного з параметрів, що визначають нормальний закон: параметр а в нормальному законі дорівнює математичному сподіванню.

Приклад 4.5. Знайдемо математичне сподівання випадкової величини ξ, рівномірно розподіленої в інтервалі (a, b). Маємо
$$\mathsf{E}\xi = \int\limits_a^b x\frac{dx}{b-a} = \frac{b^2-a^2}{2\,(b-a)} = \frac{a+b}{2}.$$
Бачимо, що математичне сподівання збігається із серединою інтервалу можливих значень випадкової величини.

Приклад 4.6. Випадкова невід'ємна величина ξ має гама розподіл з параметрами (α, β) (позначення $\xi \simeq \Gamma(\alpha, \beta)$), якщо її щільність розподілу має вигляд
$$p\,(x) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0. \end{cases}$$
Властивості гама розподілу базуються на властивостях гама-функції
$$\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx.$$
Ми скористаємося такою властивістю
$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha).$$
Знайдемо математичне сподівання випадкової величини ξ розподіленої за гама розподілом
$$\mathsf{E}\xi = \int\limits_{-\infty}^{\infty} xp(x)dx = \int\limits_0^\infty x\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx =$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} = \frac{\alpha}{\beta}.$$

Отже

$$\mathsf{E}\xi = \frac{\alpha}{\beta}.$$

Як наслідок отримаємо математичне сподівання випадкової величини ξ розподіленої за показниковим розподілом

$$p(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \ge 0. \end{cases}$$

Оскільки показниковим розподілом є частинним випадком гама розподілу з параметрами $(\alpha = 1, \beta = \lambda)$, то математичне сподівання випадкової величини ξ, розподіленої за показниковим розподілом, дорівнює

$$\mathsf{E}\xi = \frac{1}{\lambda}.$$

Приклад 4.7. Щільність двовимірної випадкової величини (ξ_1, ξ_2) задана формулою

$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x_1-a)^2}{\sigma_1^2} - 2r\frac{(x_1-a)(x_2-b)}{\sigma_1\sigma_2} + \frac{(x_2-b)^2}{\sigma_2^2}\right]}.$$

Знайдемо її математичне сподівання. Формула в цьому випадку матиме вигляд:

$$\mathsf{E}\xi_1 = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x_1\, p(x_1,\, x_2)\; dx_1\, dx_2 = \int\limits_{-\infty}^{\infty} x\, p_1(x)\, dx,$$

$$\mathsf{E}\xi_2 = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x_2\, p(x_1,\, x_2)\; dx_1\, dx_2 = \int\limits_{-\infty}^{\infty} x\, p_2(x)\, dx.$$

У прикладі попереднього розділу ми бачили, що

$$p_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma_1^2}}, \quad p_2(x) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-b)^2}{2\sigma_2^2}},$$

тому з Прикладу 4.4 цього параграфа знаходимо, що

$$\mathsf{E}\xi_1 = a, \quad \mathsf{E}\xi_2 = b.$$

Ці рівності з'ясовують імовірнісний зміст постійних a і b у нормальному законі.

4.2 Дисперсія випадкової величини

Означення 4.2. Дисперсією випадкової величини ξ називається математичне сподівання квадрата відхилення ξ від математичного сподівання $\mathsf{E}\xi$. Дисперсію позначається символом $\mathsf{D}\xi$.

Отже, за означенням

$$\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2 = \int\limits_{-\infty}^{\infty} (z - \mathsf{E}\xi)^2 dF_\xi(z). \tag{4.4}$$

Оскільки

$$(z - \mathsf{E}\xi)^2 = z^2 - 2z \cdot \mathsf{E}\xi + (\mathsf{E}\xi)^2$$

а також

$$\mathsf{E}\xi = \int\limits_{-\infty}^{\infty} z\, dF_\xi(z),$$

то формула (4.4) може бути записана інакше:

$$\mathsf{D}\xi = \int\limits_{-\infty}^{\infty} z^2 dF_\xi(z) - \left(\int\limits_{-\infty}^{\infty} z\, dF_\xi(z)\right)^2 = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2. \tag{4.5}$$

Подібно до математичного сподівання дисперсія існує не для всіх випадкових величин. Так, наприклад, розглянута нами раніше випадкова величина, розподілена за законом Коші, має нескінченну дисперсію.

Розглянемо приклади обчислення дисперсії.

Приклад 4.8. Знайдемо дисперсію випадкової величини ξ, розподіленої за законом Пуассона:

$$P\{\xi = k\} = \frac{a^k e^{-a}}{k!} \quad (k = 0, 1, 2, \ldots).$$

Обчислимо другий момент розполілу. Потім скористаємося формулою (4.5)

$$\mathsf{E}\xi^2 = \sum_{k=0}^{\infty} k^2 \frac{a^k e^{-a}}{k!} = \sum_{k=1}^{\infty} k^2 \frac{a^k e^{-a}}{k!} =$$

$$= ae^{-a} \sum_{k=1}^{\infty} \frac{(k-1+1)a^{k-1}}{(k-1)!} = a^2 e^{-a} \sum_{k=2}^{\infty} \frac{a^{k-2}}{(k-2)!} + ae^{-a} \sum_{k=1}^{\infty} \frac{a^{k-1}}{(k-1)!} =$$

$$= a^2 e^{-a} \cdot e^a + a e^{-a} \cdot e^a = a^2 + a.$$

Отже
$$\mathsf{E}\xi^2 = a^2 + a.$$

Ми знаємо, що математичне сподівання

$$\mathsf{E}\xi = a.$$

Скориставшись формулою (4.5), отримаємо

$$\mathsf{D}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = a^2 + a - a^2 = a.$$

Отже випадкова величина ξ, розподілена за законом Пуассона, має такі числові характеористики $\mathsf{E}\xi = a$, $D\xi = a$.

Приклад 4.9. Знайдемо дисперсію випадкової величини ξ, рівномірно розподіленої в інтервалі $(a,\ b)$. У даному випадку

$$\int\limits_{-\infty}^{\infty} x^2 dF_\xi\left(x\right) = \int\limits_a^b \frac{x^2}{b-a} dx = \frac{b^3 - a^3}{3\left(b-a\right)} = \frac{b^2 + ab + a^2}{3}.$$

У попередньому параграфі ми знайшли, що

$$\mathsf{E}\xi = \frac{a+b}{2}.$$

Формула (4.5) дає нам

$$\mathsf{D}\xi = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

Отже, дисперсія залежить тільки від довжини відрізка $(a,\ b)$ і є зростаючою функцією цієї довжини. Чим більший інтервал значень, яких набуває випадкова величина, тобто чим більше розсіяні її значення, тим більша дисперсія. Дисперсія, таким чином, відіграє роль міри розсіювання значень випадкової величини навколо математичного сподівання.

Приклад 4.10. Випадкова невід'ємна величина ξ має гама розподіл з параметрами (α, β) (позначення $\xi \simeq \Gamma(\alpha, \beta)$), якщо її щільність розподілу має вигляд

$$p\left(x\right) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0. \end{cases}$$

Властивості гама розподілу базуються на властивостях гама-функції

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Ми скористаємося такою властивістю гама-функції

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha).$$

Знайдемо дисперсію випадкової величини ξ розподіленої за гама розподілом. Обчислимо другий момент розполілу. Потім скористаємося формулою (4.5)

$$\mathsf{E}\xi^2 = \int_{-\infty}^{\infty} x^2 p(x) dx = \int_0^\infty x^2 \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx =$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} = \frac{(\alpha+1)\alpha}{\beta^2}.$$

Отже

$$\mathsf{E}\xi^2 = \frac{(\alpha+1)\alpha}{\beta^2}.$$

Ми знаємо, що математичне сподівання

$$\mathsf{E}\xi = \frac{\alpha}{\beta}.$$

Скориставшись формулою (4.5), отримаємо

$$\mathsf{D}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = \frac{(\alpha+1)\alpha}{\beta^2} - \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha}{\beta^2}.$$

Отже випадкова величина ξ, розподілена за гама розподілом, має такі числові характеористики

$$\mathsf{E}\xi = \frac{\alpha}{\beta},\ D\xi = \frac{\alpha}{\beta^2}.$$

Як наслідок отримаємо такі числові характеористики випадкової величини ξ розподіленої за показниковим розподілом з параметром λ:

$$\mathsf{E}\xi = \frac{1}{\lambda},\ \mathsf{D}\xi = \frac{1}{\lambda^2}.$$

Приклад 4.11. Знайдемо дисперсію випадкової величини ξ, розподіленої за нормальним законом з параметрами (a, σ^2). Щільність такого розподілу

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

Ми знаємо, що $\mathsf{E}\xi = a$, тому

$$D\xi = \int_{-\infty}^{\infty} (x-a)^2 p(x)\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2\, e^{-\frac{(x-a)^2}{2\sigma^2}} dx\,.$$

Зробимо під інтегралом заміну змінних, поклавши $z = \frac{x-a}{\sigma}$; при цьому

$$D\xi = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2\, e^{-\frac{z^2}{2}} dz\,.$$

Частинним інтегруванням знаходимо, що

$$\int_{-\infty}^{\infty} z^2\, e^{-\frac{z^2}{2}} dz = -z\, e^{-\frac{z^2}{2}}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}.$$

Звідси дістаємо, що $\mathsf{D}\xi = \sigma^2$.

Отже, ми з'ясували ймовірнісний зміст другого параметра, який визначає нормальний закон. Важливо підкреслити, що нормальний закон повністю визначається математичним сподіванням та дисперсією. Ця обставина широко використовується в застосуваннях, а також у теоретичних дослідженнях. Зауважимо, що для нормально розподіленої випадкової величини дисперсія дає змогу судити про розсіювання значень. Хоч за довільних значень дисперсії випадкова величина, розподілена за нормальним законом, може набувати всі дійсні значення, проте ймовірності значень, близьких до математичного сподівання, будуть тим більші, чим менша дисперсія. Ця обставина була нами відзначена ще в попередньому розділі при початковому ознайомленні з нормальним законом.

4.3 Дисперсія n-вимірної випадкової величини

Означення 4.3. Дисперсією (дисперсійною матрицею) n-вимірної випадкової величини $(\xi_1, \xi_2, \ldots, \xi_n)$ називається сукуність n^2 сталих b_{ij}, $1 \le i, j \le n$, визначених формулами:

$$b_{ij} = \int\limits_{-\infty}^{\infty} \dots \int\limits_{-\infty}^{\infty} (x_i - \mathsf{E}\xi_i)\,(x_j - \mathsf{E}\xi_j)\,dF\,(x_1, x_2, \dots, x_n). \qquad (4.6)$$

Оскільки за довільних дійсних $t_j, 1 \le j \le n$,

$$\int\limits_{-\infty}^{\infty} \dots \int\limits_{-\infty}^{\infty} \left(\sum_{j=1}^{n} t_j (\,x_j - \mathsf{E}\xi_j) \right)^2 dF\,(x_1, x_2, \dots, x_n) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}\, t_i\, t_j,$$

то величини b_{ij} задовольняють нерівності

$$\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}\, t_i\, t_j \ge 0, \quad \forall t_j,\, \forall t_j,\, i, j = 1, 2, \dots, n. \qquad (4.7)$$

Очевидно, що

$$b_{kk} = D\xi_k.$$

Означення 4.4. Величини b_{ij} при $i \ne j$ називаються змішаними центральними моментами 2-го порядку випадкових величин ξ_i і ξ_j. Очевидно, що $b_{ij} = b_{ji}$. Величини b_{ij} при $i \ne j$ називають ще коваріацією випадкових величин ξ_i і ξ_j і позначають $\operatorname{cov}(\xi_i, \xi_j)$. Отже

$$b_{ij} = \operatorname{cov}(\xi_i, \xi_j) = \mathsf{E}\left[(\xi_i - \mathsf{E}\xi_i)\,(\xi_j - \mathsf{E}\xi_j)\right].$$

Матрицю

$$B = \{b_{ij}\}_{i=1,\dots,n}^{j=1,\dots,n}$$

називають коваріаційною матрицею випадкового вектора $(\xi_1, \xi_2, \dots, \xi_n)$. Коваріаційна матриця додатно визначена в силу нерівностей (4.7).

Означення 4.5. Коефіцієнтом корреляції випадкових величин ξ_i та ξ_j називається величина

$$r_{ij} = \frac{\mathsf{E}\left[(\xi_i - \mathsf{E}\xi_i)\,(\xi_j - \mathsf{E}\xi_j)\right]}{\sqrt{\mathsf{E}\,(\xi_i - \mathsf{E}\xi_i)^2}\sqrt{\mathsf{E}\,(\xi_j - \mathsf{E}\xi_j)^2}} = \frac{b_{ij}}{\sqrt{b_{ii}}\sqrt{b_{jj}}}.$$

Коефіцієнт корреляції є мірою залежності між величинами ξ_i та ξ_j. Величина коефіцієнта корреляції міняється в інтервалі $[-1, 1]$. Це є наслідком нерівності Коші-Буняковського. Значення ± 1 коефіцієнт корреляції приймає лише у тому випадку, коли величини ξ_i та ξ_j звязані лінійною залежністю. Дійсно, оскільки

$$D\left(\frac{\xi_i}{\sqrt{b_{ii}}} \pm \frac{\xi_j}{\sqrt{b_{jj}}}\right) = 2(1 \pm r_{ij}) \geq 0,$$

то маємо $-1 \leq r_{ij} \leq 1$.

Рівність $r_{ij} = \pm 1$ можлива лише у тому випадку, коли

$$D\left(\frac{\xi_i}{\sqrt{b_{ii}}} \pm \frac{\xi_j}{\sqrt{b_{jj}}}\right) = 0.$$

Дисперсія дорівнює нулю лише для випадкових величин які приймають стале значення з ймовірністю 1. Тому рівність $r_{ij} = \pm 1$ можлива лише у тому випадку, коли

$$\frac{\xi_i}{\sqrt{b_{ii}}} \pm \frac{\xi_j}{\sqrt{b_{jj}}} = c.$$

Отже величини ξ_i та ξ_j звязані лінійною залежністю

$$\xi_i = \mp \frac{\sqrt{b_{ii}}}{\sqrt{b_{jj}}}\xi_j + c\sqrt{b_{ii}}.$$

Безпосередня перевірка показує, що у тому випадку, коли величини ξ_i та ξ_j звязані лінійною залежністю, то коефіцієнт корреляції приймає значення ± 1.

Для незалежних випадковмх величин, як випливає із означення, коефіцієнт корреляції дорівнює 0.

Зворотне твердження невірне. Коефіцієнт корреляції між величинами ξ та η може дорівнювати 0 для залежних випадкових величин. Як приклад, розглянемо випадкові величини ξ та η які звязані залежністю $\eta = \xi^2$, де величина ξ розподілена симетрично відносно точки $x = 0$ і має скінченний четвертий момент. Тоді $\mathsf{E}\xi = 0$, $\mathsf{E}\xi\eta = \mathsf{E}\xi^3 = 0$, $\mathsf{E}[(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)] = 0$. Отже коефіцієнт корреляції $r_{\xi\eta} = 0$ для залежних випадкових величин.

Приклад 4.12. Знайти дисперсію двовимірної випадкової величини (ξ_1, ξ_2), розподіленої за невиродженим нормальним законом $\left(r^2 \neq 1\right)$:

$$p\,(x,\,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left\{\frac{(x-a)^2}{\sigma_1^2} - 2r\frac{(x-a)(y-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{\sigma_2^2}\right\}}.$$

Згідно з результатами прикладу цього параграфа знаходимо, що

$$D\xi_1 = \sigma_1^2,\ D\xi_2 = \sigma_2^2.$$

Далі,

$$b_{12} = b_{21} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x-a)\,(y-b)\ p\,(x,y)\ dx\,dy\ =$$

$$= \int\limits_{-\infty}^{\infty} \tfrac{(y-b)}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\ e^{-\frac{(y-b)^2}{2\sigma_2^2}} dy \int\limits_{-\infty}^{\infty} (x-a)\ \ e^{-\frac{1}{2(1-r^2)}\left(\frac{x-a}{\sigma_1} - r\frac{y-b}{\sigma_2}\right)^2} dx.$$

Замінами

$$z = \frac{1}{\sqrt{1-r^2}}\left(\frac{x-a}{\sigma_1} - r\frac{y-b}{\sigma_2}\right), \qquad t = \frac{y-b}{\sigma_2}$$

зводимо вираз для b_{12} до вигляду

$$b_{12} = b_{21} = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \left(\sigma_1\sigma_2\sqrt{1-r^2}tz + r\sigma_1\sigma_2 t^2\right) e^{-\frac{t^2}{2} - \frac{z^2}{2}} dt\,dz =$$

$$= \frac{r\sigma_1\sigma_2}{2\pi} \int\limits_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt \int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz +$$

$$+ \frac{\sigma_1\sigma_2\sqrt{1-r^2}}{2\pi} \int\limits_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} dt \int\limits_{-\infty}^{\infty} z\, e^{-\frac{z^2}{2}} dz = r\sigma_1\sigma_2.$$

Отже маємо

$$b_{12} = b_{21} = r\sigma_1\sigma_2.$$

Звідси знаходимо, що

$$r = \frac{b_{12}}{\sigma_1\sigma_2} = \frac{\mathsf{E}\left[(\xi_1 - \mathsf{E}\xi_1)\,(\xi_2 - \mathsf{E}\xi_2)\right]}{\sigma_1\sigma_2}.$$

Стала r – це коефіцієнтом кореляції між випадковими величинами ξ_1 та ξ_2.

Ми бачимо, що двовимірний нормальний закон розподілу випадкового вектора $(\xi_1, \xi_2) \simeq N(a_1, a_2, \sigma_1^2, \sigma_2^2, r)$, так само як і одновимірний, цілком визначається математичним сподіванням та дисперсіями, тобто величинами $\mathsf{E}\xi_1 = a_1, \mathsf{E}\xi_2 = a_2, D\xi_1 = \sigma_1^2,\ D\xi_2 = \sigma_2^2$ та коефіцієнтом кореляції r.

Аналогічне твердження має місце і для довільного n-вимірного нормального закону.

4.4 Властивості математичного сподівання

Означення 4.6. Висловлювання щодо результату стохастичного експерименту виконується майже напевне (скорочення: м.н.), якщо ймовірність множини сприятливих елементарних подій дорівнює одиниці.

Теорема 4.1. (Центрованість.) *Якщо $\xi = 0$ м.н., то математичне сподівання $\mathsf{E}\xi = 0$.*

Теорема 4.2. (Нормованість.) *Математичне сподівання сталої C дорівнює цій сталій: $\mathsf{E}C = C$.*

Доведення. Сталу C можна розглядати як дискретну випадкову величину, яка набуватиме лише значення C з імовірністю 1. Тому $\mathsf{E}\,C = C$. □

Теорема 4.3. (Невід'ємність.) *Якщо $\xi \geq 0$ м.н., то математичне сподівання $\mathsf{E}\xi \geq 0$.*

Теорема 4.4. (Монотонність.) *Якщо $\xi \leq \eta$ м.н., то математичне сподівання $\mathsf{E}\xi \leq \mathsf{E}\eta$.*

Теорема 4.5. (Адитивність.) *Математичне сподівання суми двох випадкових величин дорівнює сумі їхніх математичних сподівань.*

$$\mathsf{E}\,(\zeta + \eta) = \mathsf{E}\zeta + \mathsf{E}\eta.$$

Доведення. Розглянемо спочатку доведення цієї теореми для випадку дискретних величин ζ та η. Нехай $a_1, a_2, \ldots, a_n, \ldots$ – можливі значення величини ζ і $p_1, p_2, \ldots, p_n, \ldots$ – їхні ймовірності; $b_1, b_2, \ldots, b_n, \ldots$ – можливі значення випадкової величини η і

$q_1, q_2, \ldots, q_n, \ldots$ – їхні ймовірності. Можливі значення випадкової величини $\zeta + \eta$ мають вигляд

$$a_n + b_k, \quad k = 1, 2, \ldots; \, n = 1, 2, \ldots$$

Позначимо через p_{nk} імовірність того, що ζ набуває значення a_n і η – значення b_k. За визначенням математичного сподівання

$$\mathsf{E}\,(\zeta + \eta) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} (a_n + b_k)\, p_{nk} =$$

$$= \sum_{n=1}^{\infty} a_n \left(\sum_{k=1}^{\infty} p_{nk} \right) + \sum_{k=1}^{\infty} b_k \left(\sum_{n=1}^{\infty} p_{nk} \right).$$

Оскільки за формулою повної ймовірності

$$\sum_{k=1}^{\infty} p_{nk} = p_n, \quad \sum_{n=1}^{\infty} p_{nk} = q_k,$$

то

$$\sum_{n=1}^{\infty} a_n \sum_{k=1}^{\infty} p_{nk} = \sum_{n=1}^{\infty} a_n p_n = \mathsf{E}\zeta,$$

$$\sum_{k=1}^{\infty} b_k \left(\sum_{n=1}^{\infty} p_{nk} \right) = \sum_{k=1}^{\infty} b_k q_k = \mathsf{E}\eta.$$

Доведення теореми для дискретних випадкових величин закінчено.

□

Наслідок 4.1. *Математичне сподівання суми скінченної кількості випадкових величин $\eta_1, \eta_2, \ldots, \eta_n$ дорівнює сумі їхніх математичних сподівань:*

$$E\,(\eta_1 + \eta_2 + \cdots + \eta_n) = E\eta_1 + E\eta_2 + \cdots + E\,\eta_n.$$

Доведення. За тільки що доведеною теоремою

$$\mathsf{E}\,(\eta_1 + \eta_2 + \cdots + \eta_n) = \mathsf{E}\,\eta_1 + \mathsf{E}\,(\eta_2 + \cdots + \eta_n) =$$

$$= \mathsf{E}\,\eta_1 + \mathsf{E}\,\eta_2 + \mathsf{E}\,(\eta_3 + \cdots + \eta_n) = \cdots =$$

$$= \mathsf{E}\,\eta_1 + \mathsf{E}\,\eta_2 + \cdots + \mathsf{E}\,\eta_n.$$

□

Наслідок 4.2. *Математичне сподівання суми*

$$S_\mu = \eta_1 + \cdots + \eta_\mu,$$

де μ – випадкова величина, що приймає лише цілочисельні додатні значення, випадкові величини $\eta_1, \eta_2, \ldots$ не залежать від μ, математичне сподівання μ скінченне і ряд

$$\sum_{i=1}^{\infty} \mathsf{E}|\eta_j| \cdot P\{\mu \geq j\}$$

збігається, скінченне і дорівнює

$$\mathsf{E}\, S_\mu = \sum_{i=1}^{\infty} \mathsf{E}\,\eta_j \cdot P\{\mu \geq j\}.$$

Доведення. Справді, умовне математичне сподівання S_μ за умови, що $\mu = k$, дорівнює

$$\mathsf{E}\{S_\mu\,|\mu = k\} = \mathsf{E}\,\eta_1 + \cdots + \mathsf{E}\eta_k.$$

Безумовне математичне сподівання

$$\mathsf{E}\, S_\mu = \sum_{k=1}^{\infty} \mathsf{E}\,(S_\mu\,|\mu = k)\, P\{\mu = k\} = \sum_{k=1}^{\infty} P\{\mu = k\} \cdot \sum_{j=1}^{k} \mathsf{E}\eta_j =$$

$$= \sum_{j=1}^{\infty} \mathsf{E}\,\eta_j \cdot \sum_{k=j}^{\infty} P\{\mu = k\} = \sum_{j=1}^{\infty} \mathsf{E}\,\eta_j \cdot P\{\mu \geq j\}.$$

Якщо доданки $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ однаково розподілені, тобто

$$P\{\eta_1 < x\} = P\{\eta_2 < x\} = \cdots = F(x)\,,$$

то

$$\mathsf{E}\, S_\mu = \mathsf{E}\eta_1 \cdot \mathsf{E}\mu.$$

Справді,

$$\mathsf{E}S_\mu = \sum_{k=1}^{\infty} P\{\mu = k\} \sum_{j=1}^{k} \mathsf{E}\eta_j = \mathsf{E}\eta_1 \sum_{k=1}^{\infty} k \cdot P\{\mu = k\} = \mathsf{E}\eta_1 \cdot \mathsf{E}\mu.$$

□

Приклад 4.13. Кількість космічних частинок, що потрапляють на дану ділянку за одиницю часу, є випадкова величина μ, яка підпорядкована закону Пуассона з параметром a. Кожна частинка несе енергію ε. Знайти середню енергію S, яку дістає за одиницю часу ділянка.

Згідно з наслідком 4.2 маємо:

$$\mathsf{E}\,(S) = \mathsf{E}\,\varepsilon \cdot \mathsf{E}\,\mu = a\mathsf{E}\,\varepsilon.$$

Приклад 4.14. Стрільба провадиться до n-го влучання у мішень. Вважаючи, що постріли відбуваються незалежно один від одного і ймовірність влучання при кожному пострілі дорівнює p, знайти математичне сподівання витрати снарядів до n-го влучання.

Позначимо через μ_k кількість снарядів, витрачених від $(k-1)$-го до k-го влучання. Очевидно, що витрата снарядів на n влучань дорівнює $\mu = \mu_1 + \mu_2 + \ldots + \mu_n$. За щойно доведеною теоремою

$$\mathsf{E}\,\mu = \mathsf{E}\,\mu_1 + \mathsf{E}\,\mu_2 + \cdots + \mathsf{E}\,\mu_n.$$

Але

$$\mathsf{E}\,\mu_1 = \mathsf{E}\,\mu_2 = \ldots = \mathsf{E}\,\mu_n.$$

Тому

$$\mathsf{E}\mu = n\mathsf{E}\mu_1.$$

Нам залишається знайти $\mathsf{E}\mu_1$. Для цього знайдемо ймовірність P_m того, що вперше в ціль влучить m-й снаряд. Інакше кажучи, P_m є ймовірність того, що при перших $m-1$ пострілах будуть промахи і лише при m-у пострілі влучать у ціль. Таким чином,

$$P_m = q^{m-1}p, \quad q = 1-p.$$

Тепер

$$\mathsf{E}\mu_1 = \sum_{m=1}^{\infty} mpq^{m-1} = p\frac{d}{dq}\left(\sum_{m=0}^{\infty} q^m\right) = p\cdot\frac{d}{dq}\left(\frac{1}{1-q}\right) = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

Отже

$$\mathsf{E}\,\mu = \frac{n}{p}.$$

Теорема 4.6. *Математичне сподівання добутку незалежних випадкових величин ξ та η дорівнює добуткові їхніх математичних сподівань.*

Доведення. Якщо величини ξ та η дискретні, $a_1, a_2, \ldots, a_k, \ldots$ можливі значення ξ та $p_1, p_2, \ldots, p_k, \ldots$– їхні ймовірності; $b_1, b_2, \ldots, b_n, \ldots$ – можливі значення η і $q_1, q_2, \ldots, q_n, \ldots$ – їхні ймовірності, то ймовірність того, що ξ набуває значення a_k, а η – значення b_n, дорівнює $p_k q_n$. За означенням математичного сподівання

$$\mathsf{E}\,\xi{\cdot}\eta = \sum_{k,n} a_k b_n p_k q_n = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} a_k b_n p_k q_n = \sum_{k=1}^{\infty} a_k p_k \sum_{n=1}^{\infty} b_n q_n = \mathsf{E}\,\xi{\cdot}\mathsf{E}\,\eta.$$

□

Наслідок 4.3. (Однорідність.) *Сталий множник можна виносити за знак математичного сподівання:*

$$E(C \cdot \xi) = C \cdot E(\xi).$$

Це твердження очевидне через те, що, яке б не було ξ, сталу C і ξ можна розглядати як незалежні випадкові величини.

4.5 Властивості дисперсії

За умови квадратичної інтегровності випадкової величини дисперсія має такі властивості.

Теорема 4.7. *Дисперсія сталої дорівнює нулю.*

Доведення. Згідно з теоремою 1

$$\mathsf{D}\xi = \mathsf{E}(C - \mathsf{E}C)^2 = \mathsf{E}(C - C)^2 = 0.$$

□

Теорема 4.8. *Дисперсія $D\xi \geq 0$.*

Доведення. Випливає з позитивності математичного сподівання, оскільки величина $(\xi - \mathsf{E}\xi)^2 \geq 0$. □

Теорема 4.9. *Дисперсія $D\xi = 0$ тоді й тільки тоді, коли $\xi = E\xi$ майже напевне,*

Доведення. Виводиться з невід'ємності квадратичної функції та з теореми про властивості математичного сподівання, а саме, з властивостей невід'ємності і додатності. □

Теорема 4.10. *Дисперсію* $D\xi$ *можна обчислювати за формулою* $D\xi = E\xi^2 - (E\xi)^2$.

Доведення. Для доведення обчислимо за лінійністю математичного сподівання

$$\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2 = \mathsf{E}(\xi^2 - 2\xi\mathsf{E}\xi + (\mathsf{E}\xi)^2) = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2.$$

□

Теорема 4.11. *Якщо* a, b *– сталі, то*

$$D(a\xi + b) = a^2 D\xi.$$

Доведення. За означенням дисперсії та лінійністю математичного сподівання

$$\mathsf{D}(a\xi + b) = \mathsf{E}((a\xi + b) - \mathsf{E}(a\xi + b))^2 = \mathsf{E}(a\xi - a\mathsf{E}\xi)^2 = a^2\mathsf{D}\xi.$$

□

Теорема 4.12. *Математичне сподівання і дисперсія пов'язані таким співвідношенням*

$$D\xi = \min_{c\in\mathbb{R}} E(\xi - c)^2, \quad E\xi = \arg\min_{c\in\mathbb{R}} E(\xi - c)^2.$$

Доведення. Для доведення обчислимо

$$\mathsf{E}(\xi - c)^2 = \mathsf{E}(\xi - \mathsf{E}\xi + \mathsf{E}\xi - c)^2 = \mathsf{E}(\xi - \mathsf{E}\xi)^2 + (\mathsf{E}\xi - c)^2 + 2\mathsf{E}[(\xi - \mathsf{E}\xi)(\mathsf{E}\xi - c)] =$$

$$= \mathsf{E}(\xi - \mathsf{E}\xi)^2 + (\mathsf{E}\xi - c)^2 \geq \mathsf{E}(\xi - \mathsf{E}\xi)^2 = \mathsf{D}\xi.$$

причому рівність має місце тоді й тільки тоді, коли $c = \mathsf{E}\xi$. □

Теорема 4.13. *Дисперсія суми незалежних випадкових величин* ξ *та* η *дорівнює сумі їхніх дисперсій:*

$$D\,(\xi + \eta) = D\xi + D\eta.$$

Доведення. Справді,

$$\mathsf{D}\,(\xi + \eta) = \mathsf{E}[\xi + \eta - \mathsf{E}\,(\xi + \eta)]^2 = \mathsf{E}[(\xi - \mathsf{E}\xi) + (\eta - \mathsf{E}\eta)]^2 =$$

$$= D\xi + D\eta + 2\mathsf{E}\left[(\xi - \mathsf{E}\xi)\,(\eta - \mathsf{E}\eta)\right].$$

Величини ξ та η незалежні, тому незалежні також величини $\xi - \mathsf{E}\xi$ та $\eta - \mathsf{E}\eta$, звідки

$$\mathsf{E}\left[(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)\right] = \mathsf{E}(\xi - \mathsf{E}\xi) \cdot \mathsf{E}(\eta - \mathsf{E}\eta) = 0.$$

Твердження доведене. $\square$

Наслідок 4.4. *Якщо $\xi_1, \xi_2, \ldots, \xi_n$ – випадкові величини, кожна з яких незалежна від суми попередніх, тобто ξ_k – незалежна від $\xi_1 + \xi_2 + \cdots + \xi_{k-1}$ при $k = 2, 3, \ldots, n$, то*

$$D(\xi_1 + \xi_2 + \ldots + \xi_n) = D\xi_1 + D\xi_2 + \ldots + D\xi_n.$$

Доведення. Справді,

$$\mathsf{D}(\xi_1 + \cdots + \xi_n) = D\left[(\xi_1 + \cdots + \xi_{n-1}) + \xi_n\right] =$$
$$= \mathsf{D}(\xi_1 + \cdots + \xi_{n-1}) + D\xi_n =$$
$$= \mathsf{D}(\xi_1 + \cdots + \xi_{n-2}) + \mathsf{D}\xi_{n-1} + \mathsf{D}\xi_n = \cdots =$$
$$= \mathsf{D}\xi_1 + \mathsf{D}\xi_2 + \cdots + \mathsf{D}\xi_n.$$

$\square$

Наслідок 4.5. *Дисперсія суми скінченної кількості попарно незалежних випадкових величин $\xi_1, \xi_2, \ldots, \xi_n$ дорівнює сумі їхніх дисперсій.*

Доведення. Справді,

$$\mathsf{D}(\xi_1 + \ldots + \xi_n) = \mathsf{E}\left[\sum_{k=1}^{n} (\xi_k - \mathsf{E}\xi_k)\right]^2 =$$

$$= \mathsf{E}\sum_{j=1}^{n}\sum_{k=1}^{n} (\xi_k - \mathsf{E}\xi_k)(\xi_j - \mathsf{E}\xi_j) =$$

$$= \sum_{k=1}^{n}\sum_{j=1}^{n} \mathsf{E}\left[(\xi_k - \mathsf{E}\xi_k)(\xi_j - \mathsf{E}\xi_j)\right] =$$

$$= \sum_{k=1}^{n} \mathsf{D}\xi_k + \sum_{k \neq j} \mathsf{E}\left[(\xi_k - \mathsf{E}\xi_k)(\xi_j - \mathsf{E}\xi_j)\right].$$

Із незалежності довільної пари випадкових величин ξ_k та ξ_j $(k \neq j)$ випливає, що при $k \neq j$

$$\mathsf{E}\left[(\xi_k - \mathsf{E}\xi_k)(\xi_j - \mathsf{E}\xi_j)\right] = 0.$$

Ця рівність завершує доведення твердження. □

Наслідок 4.6. *Нормованим відхиленням випадкової величини називаєтъся відношення*

$$\frac{\xi - E\xi}{\sqrt{D\xi}}.$$

Доведемо, що

$$D\left(\frac{\xi - E\xi}{\sqrt{D\xi}}\right) = 1.$$

Доведення. Справді, ξ та $\mathsf{E}\,\xi$ як випадкові величини незалежні, тому за теоремами 4.7, 4.11 та 4.13

$$D\left(\frac{\xi - \mathsf{E}\,\xi}{\sqrt{D\xi}}\right) = \frac{D\xi + D\left(-\mathsf{E}\xi\right)}{D\xi} = \frac{D\xi}{D\xi} = 1.$$

□

Наслідок 4.7. *Якщо ξ та η – незалежні випадкові величини, то*

$$D(\xi - \eta) = D\xi + D\eta.$$

Доведення. Справді за теоремами 4.11 та 4.13

$$\mathsf{D}\left(\xi - \eta\right) = D\left[\xi + (-1)\,\eta\right] = \mathsf{D}\xi + (-1)^2\mathsf{D}\eta = D\xi + D\eta.$$

□

Приклад 4.15. Теореми 4.5 та 4.13 можна застосувати при обчисленні математичного сподівання та дисперсії числа μ появ події A при n незалежних випробовуваннях. Якщо p_k є ймовірність події A при k-му випробовуванні, то має місце так звана схема Пуассона. Позначимо через μ_k кількість появ події А при k-му випробовуванні. Очевидно, що μ_k є випадкова величина, яка набуває значення 0 та 1 з імовірностями $q_k = 1 - p_k$ та p_k відповідно. Величина μ, таким чином, може бути подана як сума

$$\mu = \mu_1 + \mu_2 + \ldots + \mu_n.$$

Оскільки
$$\mathsf{E}\mu_k = 0 \cdot q_k + 1 \cdot p_k = p_k,$$
$$D\mu_k = \mathsf{E}\mu_k^2 - (\mathsf{E}\mu_k)^2 = 0 \cdot q_k + 1 \cdot p_k - p_k^2 = p_k\,(1 - p_k) = p_k q_k,$$
то доведені теореми дозволяють нам зробити висновок, що для схеми Пуассона
$$\mathsf{E}\mu = p_1 + p_2 + \ldots + p_n,$$
$$D\mu = p_1 q_1 + p_2 q_2 + \ldots + p_n q_n.$$
Для випадку схеми Бернуллі $p_k = p$ і, отже,
$$\mathsf{E}\mu = np, \quad D\mu = npq.$$
Звідси знаходимо, що
$$\mathsf{E}\frac{\mu}{n} = p, \;\; D\frac{\mu}{n} = \frac{pq}{n}.$$

4.6 Моменти випадкової величини

Моментом k-го порядку випадкової величини ξ називається математичне сподівання величини $(\xi - a)^k$:
$$\nu_k\,(a) = \mathsf{E}(\xi - a)^k. \tag{4.8}$$
Якщо $a = 0$, то момент називається **початковим**. Легко бачити, що початковий момент першого порядку є не що інше, як математичне сподівання випадкової величини ξ.

Якщо $a = \mathsf{E}\xi$, то момент називається **центральним.** Центральний момент першого порядку довільної випадкової величини ξ дорівнює нулю, а центральний момент другого порядку є не що інше, як дисперсія цієї величини.

Початкові моменти будемо далі позначати символом ν_k без аргументу а у формулі (4.8). Центральні моменти позначатимемо буквою μ_k, указуючи в обох випадках нижнім індексом порядок моменту. Між центральним і початковим моментами існує простий зв'язок. Справді,
$$\mu_n = \mathsf{E}(\xi - \mathsf{E}\xi)^n = \sum_{k=0}^{n} C_n^k \mathsf{E}\xi^k (-\mathsf{E}\xi)^{n-k} = \sum_{k=0}^{n} (-1)^{n-k} C_n^k \nu_k \nu_1^{n-k}. \tag{4.9}$$

Через те, що $\nu_1 = \mathsf{E}\xi$ та $\nu_0 = 1$, маємо

$$\mu_n = \sum_{k=2}^{n} (-1)^{n-k} C_n^k \nu_k \nu_1^{n-k} + (-1)^{n-1} (n-1) \nu_1^n. \tag{4.10}$$

Для $n = 2$ ця формула була виведена нами при визначенні поняття дисперсії. Випишемо цю залежність для перших значень n.

$$\begin{gathered} \mu_0 = 1, \\ \mu_1 = 0, \\ \mu_2 = \nu_2 - \nu_1^2, \\ \mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3, \\ \mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4. \end{gathered}$$

Ці перші моменти відіграють особливо важливу роль у математичній статистиці.

Величина

$$m_k = \mathsf{E}|\xi - \mathsf{E}\xi|^k \tag{4.11}$$

називається **абсолютним моментом k-го порядку.**

Математичне сподівання $\mathsf{E}(\xi - a)^k$ обчислюється за формулою

$$\nu_k(a) = \int_{-\infty}^{\infty} (x - a)^k \, dF(x), \tag{4.12}$$

де $F(x)$ є функція розподілу величини ξ. Це твердження є частинним випадком такого твердження.

Теорема 4.14. *Якщо $F(x)$ функція розподілу випадкової величини ξ, а $f(x)$ – неперервна функція, то математичне сподівання випадкової величини $\eta = f(\xi)$ обчислюється за формулою*

$$\mathsf{E}f(\xi) = \int_{-\infty}^{\infty} f(x) \, dF(x).$$

Аналогічне твердження має місце і для випадкової величини $\eta = f(\xi_1, \xi_2, \ldots, \xi_n)$, яка є неперервною функцією від n випадкових величин $\xi_1, \xi_2, \ldots, \xi_n$. А саме:

Теорема 4.15. *Математичне сподівання випадкової величини* $\eta = f(\xi_1, \xi_2, \ldots, \xi_n)$*, яка є неперервною функцією від* n *випадкових величин* $\xi_1, \xi_2, \ldots, \xi_n$ *обчислюється за формулою*

$$\mathsf{E}\eta = \mathsf{E}f(\xi_1, \xi_2, \ldots, \xi_n) = \int\limits_{-\infty}^{\infty} x\, dF_\eta(x) =$$
$$= \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_n)\, dF(x_1, \ldots, x_n),$$

де $F(x_1, x_2, \ldots, x_n)$ *є функція розподілу* n*-вимірної величини* $(\xi_1, \xi_2, \ldots, \xi_n)$.

Оскільки випадкова величина ξ має математичне сподівання тільки тоді, коли інтеграл, яким воно виражається, абсолютно збігається, то, очевидно, що момент k-го порядку у випадкової величини існує тільки тоді, коли збігається інтеграл

$$\int\limits_{-\infty}^{\infty} |x|^k\, dF_\xi(x).$$

Із цього зауваження випливає, що коли випадкова величина ξ має момент k-го порядку, то вона має також моменти всіх додатних порядків, менших ніж k.

Справді, через те що $|x|^k > |x|^r$ при $r < k$, коли $|x| > 1$, то

$$\int\limits_{-\infty}^{\infty} |x|^r\, dF_\xi(x) = \int\limits_{|x|\le 1} |x|^r\, dF_\xi(x) + \int\limits_{|x|>1} |x|^r\, dF_\xi(x) \le$$
$$\le \int\limits_{|x|\le 1} |x|^r\, dF_\xi(x) + \int\limits_{|x|>1} |x|^k\, dF_\xi(x).$$

Перший інтеграл у правій частині нерівності збігається внаслідок скінченності границь, другий інтеграл збігається за припущенням. Це доведодить наше твердження.

Приклад 4.16. Знайдемо моменти випадкової величини ξ розподіленої за гама розподілом з параметрами (α, β) (позначення $\xi \simeq \Gamma(\alpha, \beta)$). Щільність розподілу має вигляд

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \ge 0. \end{cases}$$

Для такого розподілу

$$\mathsf{E}\xi^k = \int_{-\infty}^{\infty} x^k p(x)dx = \int_0^{\infty} x^k \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx =$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}} = \frac{(\alpha+k-1)\cdot(\alpha+k-2)\cdots\alpha}{\beta^k}.$$

Тому

$$\mathsf{E}\xi^k = \frac{(\alpha+k-1)\cdot(\alpha+k-2)\cdots\alpha}{\beta^k}, \quad k = 1, 2, \ldots$$

Отже випадкова величина ξ, розподілена за гама розподілом, має такі моменти

$$\mathsf{E}\xi = \frac{\alpha}{\beta},\ \ \mathsf{E}\xi^2 = \frac{(\alpha+1)\alpha}{\beta^2},\ \ \mathsf{E}\xi^3 = \frac{(\alpha+2)(\alpha+1)\alpha}{\beta^3}, \ldots$$

$$\mathsf{E}\xi^k = \frac{(\alpha+k-1)\cdot(\alpha+k-2)\cdots\alpha}{\beta^k}, \ldots.$$

Як наслідок отримаємо такі моменти випадкової величини ξ розподіленої за показниковим розподілом з параметром λ:

$$\mathsf{E}\xi = \frac{1}{\lambda},\ \ \mathsf{E}\xi^2 = \frac{2}{\lambda^2},\ \ \mathsf{E}\xi^3 = \frac{6}{\lambda^3}, \ldots, \mathsf{E}\xi^k = \frac{k!}{\lambda^k}, \ldots$$

Приклад 4.17. Знайдемо центральні й абсолютні центральні моменти випадкової величини, розподіленої за нормальним законом.

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

Маємо

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

При непарному k унаслідок непарності підінтегральної функції $\mu_k = 0$. При k парному

$$\mu_k = m_k = \sqrt{\frac{2}{\pi}} \sigma^k \int_0^{\infty} x^k e^{-\frac{x^2}{2}} dx.$$

Заміною $x^2 = 2z$ зводимо цей інтеграл до вигляду

$$\mu_k = m_k = \sqrt{\frac{2}{\pi}}\sigma^k 2^{\frac{k-1}{2}} \int\limits_0^\infty z^{\frac{k-1}{2}} e^{-z} dz =$$

$$= \sqrt{\frac{2}{\pi}}\sigma^k 2^{\frac{k-1}{2}} \Gamma\left(\frac{k+1}{2}\right) = \frac{k\,!}{2^{\frac{k}{2}} \cdot \left(\frac{k}{2}\right)!}\sigma^k.$$

При k непарному

$$m_k = \sqrt{\frac{2}{\pi}}\sigma^k \int\limits_0^\infty x^k e^{-\frac{x^2}{2}} dx =$$

$$= \frac{\sigma^k}{\sqrt{\pi}} 2^{\frac{k}{2}} \Gamma\left(\frac{k+1}{2}\right) = \sqrt{\frac{2^k}{\pi}} \left(\frac{k-1}{2}\right)!\,\sigma^k.$$

Моменти функції розподілу не можуть бути довільними величинами. Через те, що які б не були сталі $t_0, t_1, \ldots, t_n$, квадратична форма

$$I_n = \int\limits_{-\infty}^{\infty} \left(\sum_{k=0}^{n} t_k (x-a)^k\right)^2 dF\left(x\right) = \sum_{k=0}^{n}\sum_{j=0}^{n} \nu_{k+j}\left(a\right) t_k t_j \geq 0$$

невід'ємна, із теорії квадратичних форм відомо, що матриця

$$\begin{vmatrix} \nu_0\left(a\right) & \nu_1\left(a\right) & \cdots & \nu_k\left(a\right) \\ \nu_1\left(a\right) & \nu_2\left(a\right) & \cdots & \nu_{k+1}\left(a\right) \\ \vdots & \vdots & \ddots & \vdots \\ \nu_k\left(a\right) & \nu_{k+1}\left(a\right) & ... & \nu_{2k}\left(a\right) \end{vmatrix}$$

розмірності $(k+1) \times (k+1)$ невід'ємно визначена для всіх $k = 0, 1, 2, \ldots, n$.

Аналогічні нерівності підпорядковані також абсолютні моменти.

У прикладах із попереднього параграфа видно, що два перші моменти випадкової величини повністю визначають її розподіл (нормальний, Пуассона, рівномірний, Сімпсона та ін.). У математичній статистиці значну роль відіграють закони розподілу, які залежать від більшої, ніж два, кількості параметрів. Якщо наперед відомо, що випадкова величина підпорядкована цілком певному закону, а невідомі лише значення параметрів, то ці невідомі параметри в найважливіших випадках визначаються через перші моменти випадкової величини. Якщо ж нам невідомо, до якого виду належить функція розподілу випадкової величини ξ, то, взагалі кажучи, не тільки знання перших, але й знання всіх цілих моментів не дає можливості визначити невідому функцію розподілу $F_\xi(x)$. Виявляється, що можна побудувати приклади різних функцій розподілу з однаковими моментами всіх цілих порядків. У зв'язку із цим виникає задача (проблема моментів): дана послідовність дійсних чисел

$$c_0, c_1, c_2, \ldots, c_n, \ldots$$

За яких умов існує одна й тільки одна функція розподілу $F(x)$, для якої при всіх n мають місце рівності

$$c_n = \int\limits_{-\infty}^{\infty} x^n dF(x).$$

Ця задача повністю розв'язана. Ми не будемо розглядати її у цій книзі.

Серед інших числових характеристик випадкових величин найбільш істотну роль відіграють так звані семіінваріанти. Розглянемо визначення семіінваріантів у наступних розділах. Тут відзначимо тільки таке. При додаванні незалежних випадкових величин момент суми не дорівнює сумі моментів доданків. Для моменту суми незалежних доданків ξ та η має місце така формула:

$$\mathsf{E}(\xi+\eta)^n = \sum_{k=0}^{n} C_n^k \, \mathsf{E}\xi^k \cdot \mathsf{E}\eta^{n-k}.$$

Семіінваріанти різних порядків мають ту властивість, що семіінваріант суми незалежних доданків дорівнює сумі їхніх семіінваріантів. Виявляється, що семіінваріант довільного порядку k є раціональна функція моментів порядків менших або рівних k.

Медіаною розподілу $F(x)$ називається таке значення аргументу m для якого виконуються нерівності

$$F(m) \le \frac{1}{2} \le F(m+0).$$

Якщо функція $F(x)$ неперервна, то існує принамі одне значення m, для якого виконується рівність

$$F(m) = \frac{1}{2}.$$

Якщо крива $y = F(x)$ та пряма $y = \frac{1}{2}$ мають спільний замкнутий відрізок $[a, b]$, то кожна точка цього відрізкв є медіаною.

Медіана існує для кожного розподілу, на відміну від математичного сподівання, яке може не існувати.

Медіана нормального розподілу дорівнює його математичному сподіванню.

Медіана розподілу має таку властивість.

Теорема 4.16. *Абсолютний момент* $\mathsf{E}|\xi - c|$ *для неперервного розподілу* $F(x)$ *в залежності від значення* c *приймає мінімальне значення, якщо* c *дорівнює медіані розподілу.*

Доведення. Доведення випливає із такої рівності

$$\mathsf{E}|\xi - c| = \begin{cases} \mathsf{E}|\xi - m| + 2\int_m^c (c - x)dF(x), & c > m, \\ \mathsf{E}|\xi - m| + 2\int_c^m (x - c)dF(x), & c < m. \end{cases}$$

□

Точно так як була визначена медіана визначається **квантиль розподілу** порядку p, $0 < p < 1$, це таке значення аргументу x_p, для якого виконуються нерівності

$$F(x_p) \leq p \leq F(x_p + 0).$$

Для неперервного розподілу квантиль порядку p називають корінь рівняння $F(x_p) = p$. Квантиль порядку $p = 1/2$ є медіаною розподілу.

Якщо функція розподілу неперервна, тобто має щільність розподілу, то **модою розподілу** називається значення аргументу при якому щільність досягає найбільшого значення.

Для нормального розподілу мода співпадає з медіаною та з математичним сподіванням.

5 Закон великих чисел

Величезний досвід, набутий людством, навчає нас, що явища, які мають імовірність дуже близьку до одиниці, майже обов'язково відбуваються. Так само події, імовірність появи яких дуже мала (іншими словами, дуже близька до нуля), настають дуже рідко. Ця обставина відіграє важливу роль для всіх практичних висновків теорії ймовірностей, оскільки досвід дає нам право в практичній діяльності вважати малоймовірні події практично неможливими, а події, що відбуваються з імовірністю, близькою до одиниці, практично достовірними. При цьому на цілком природне запитання – якою має бути ймовірність, щоб ми могли вважати подію практично неможливою – однозначної відповіді дати не можна. І це цілком зрозуміло, так як у практичній діяльності зазвичай необхідно брати до уваги властивості тих подій, з якими доводиться мати справу. Так, наприклад, якщо при вимірюванні відстані між двома селищами з'ясувалося, що вона дорівнює 5320 м, причому абсолютна похибка цього виміру з імовірністю 0,02 перевищує 10 м, то можна знехтувати можливістю такої похибки і вважати, що віддаль насправді дорівнює 5320 м. Отже, у цьому прикладі подію з імовірністю 0,02 вважаємо практично неможливою й у своїй практичній діяльності її не беремо до уваги. У той же час в інших випадках нехтувати ймовірністю 0,02 і навіть ще меншими ймовірностями не можна. Так, якщо при будівництві великої гідроелектростанції, яке потребує великих матеріальних витрат, з'ясовується, що ймовірність катастрофічного паводку в даних умовах дорівнює 0,02, то ця ймовірність буде вважатися великою, і при проектуванні станції така подія повинна бути врахована, а не відкинута, як це було зроблено в попередньому прикладі. Таким чином, лише вимоги практики можуть підказати нам критерії, згідно з якими ми будемо вважати ті або інші події практично неможливими або практично достовірними. Однак слід зауважити, що довільна подія, якою б малою не була її ймовірність, може відбутися, і якщо число випробовувань, у кожному з яких вона може з'явитися з однією і тією ж імовірністю, дуже велике, то ймовірність хоча б одноразової її появи може стати як завгодно близькою до одиниці. Цю обставину завжди треба мати на увазі. Але якщо ймовірність якоїсь події дуже мала, то важко сподіватися на її появу в якому-небудь наперед визначеному випробовуванні. Так, якщо хтось твердить, що при першій же роздачі колоди карт між чотирма партнерами кожний отримає карти лише однієї масті, то природно запідозрити, що при роздачі користувались якимсь певним правилом, наприклад, розмістили карти в порядку, наперед відомому тому, хто їх роздає. Таке припущення ґрунтується на тому, що ймовірність подібного тасування

$$(9!)^{44!}/_{36!} < 1,1 \cdot 10^{-18},$$

тобто дуже мала. Проте не можна твердити, що такий розклад не може відбутися. І справді, одного разу факт такого розкладу карт був зареєстрований. Цей приклад досить добре ілюструє різницю між поняттям практичної неможливості й неможливості, так би мовити, категоричної. Із попереднього ясно, що в практичній діяльності особливо велике значення мають події з імовірностями, близькими до одиниці й до нуля. Тому ясно, що однією з найважливіших задач теорії ймовірностей має бути встановлення подій, імовірності яких близькі до цих крайніх можливих значень. Закон великих чисел є одним із таких найважливіших тверджень теорії ймовірностей.

Під законом великих чисел тепер було б природно розуміти всю сукупність тверджень, в яких визначено з імовірністю, як завгодно близькою до одиниці, що настане певна подія, яка залежить від необмежено зростаючого числа випадкових подій, кожна з яких має на неї незначний вплив.

Це загальне уявлення про теореми типу закону великих чисел можна сформулювати і дещо інакше.

Нехай маємо послідовність випадкових чисел

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

Розглянемо величини ζ_n, що є деякими заданими симетричними функціями від перших n величин послідовності $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, тобто

$$\zeta_n = f_n(\xi_1, \xi_2, \ldots, \xi_n).$$

Якщо існує така послідовність сталих $a_1, a_2, \ldots, a_2, \ldots$, що при довільному $\varepsilon > 0$

$$\lim_{n \to \infty} P\left\{|\zeta_n - a_n| < \varepsilon\right\} = 1, \tag{5.1}$$

то послідовність $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ підпорядкована закону великих чисел із заданими функціями $\zeta_n = f_n(\xi_1, \xi_2, \ldots, \xi_n)$.

Зазвичай у поняття закону великих чисел вкладається певний, цілком визначений зміст. А саме, обмежуються тим випадком, коли ζ_n є середнє арифметичне величин $\xi_1, \xi_2, \ldots, \xi_n$.

Якщо у співвідношенні (5.1) всі величини a_n дорівнюють одній і тій же величині a, то кажуть, що величини ζ_n збігаються за ймовірністю до a. Співвідношення (5.1) в цій термінології означає, що $\zeta_n - a_n$ збігається за ймовірністю до нуля.

Практичне значення закону великих чисел можна проілюструвати на такому прикладі. За сучасними фізичними поглядами будь-який газ складається з величезної кількості окремих частинок, що перебувають у безперервному хаотичному русі. Про кожну окрему молекулу теорія ймовірностей не може сказати наперед, з якою швидкістю вона буде рухатись і в якому місці буде перебувати в кожний заданий момент часу. Теорія ймовірностей може лише обчислити за даних умов, яка частина всіх молекул буде рухатися із заданою швидкістю, або яка частина з них буде міститися в даному об'ємі. Але, власне, саме це треба знати фізику, оскільки основні характеристики стану газу – тиск, температура, густина – визначаються не поведінкою будь-якої однієї молекули, а їхньою спільною взаємодією. Так, наприклад, тиск газу визначається сумарним впливом молекул, які вдарилися об пластинку одиничної площі за одиницю часу. Число ударів і швидкості молекул, що ударяються, змінюються залежно від випадку, проте за законом великих чисел (ми його доведемо в наступному параграфі в формі Чебишева) тиск має бути майже сталим. Цей нівелювальний вплив закону великих чисел у фізичних явищах завдяки дуже великій кількості інгредієнтів, які беруть участь у явищах, виявляється з колосальною точністю. Досить згадати, що, наприклад, у звичайних умовах навіть дуже точними експериментами важко встановити відхилення від закону Паскаля, відомого кожному школяреві. Цей надзвичайно точний збіг з досвідом навіть служив, як це не дивно, аргументом противникам молекулярної будови матерії: коли б матерія мала молекулярну будову, то спостерігалися б і відхилення від закону Паскаля. Ці відхилення, так звані флуктуації тиску, справді вдалося спостерігати, коли навчилися відокремлювати порівняно невеликі кількості молекул, у результаті чого вплив окремих молекул ще не повністю нівелювався і був досить сильним. Наявність же таких флуктуацій за самим змістом теореми Чебишева є очевидним фактом.

5.1 Закон великих чисел у формі чебишева

Означення 5.1. Для послідовності випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ виконується закон великих чисел, якщо різниця між середнім арифметичним випадкових величин та середнім арифметичним математисних сподівань цих випадкових величин задовольняє граничне співвідношення

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}\xi_k\right| < \varepsilon\right\} = 1.$$

Доведення наступних теорем проведемо досить простим і в той же час дуже сильним методом, указаним Чебишевим, який базується на використанні елементарної нерівності.

Лема 5.1. (Нерівність Чебишева.) *Для довільної випадкової величини, що має скінченну дисперсію, при довільному $\varepsilon > 0$ має місце нерівність:*

$$P\left\{|\xi - E\xi| \geq \varepsilon\right\} \leq \frac{D\xi}{\varepsilon^2}.$$

Доведення. Якщо позначити через $F(x)$ функцію розподілу величини ξ, то можна записати, що

$$P\left\{|\xi - \mathsf{E}\xi| \geq \varepsilon\right\} = \int\limits_{|x-\mathsf{E}\xi|\geq\varepsilon} dF(x).$$

Оскільки в області інтегрування $|x - \mathsf{E}\xi| \geq \varepsilon$, то ясно, що

$$\int\limits_{|x-\mathsf{E}\xi|\geq\varepsilon} dF(x) \leq \int\limits_{|x-\mathsf{E}\xi|\geq\varepsilon} \frac{(x-\mathsf{E}\xi)^2}{\varepsilon^2} dF(x)$$

Ми лише посилимо цю нерівність, поширивши інтегрування на всі значення x:

$$\int\limits_{|x-\mathsf{E}\xi|\geq\varepsilon} dF(x) \leq \int\limits_{-\infty}^{\infty} \frac{(x-\mathsf{E}\xi)^2}{\varepsilon^2} dF(x) = \frac{D\xi}{\varepsilon^2}.$$

Нерівність Чебишева доведена. □

Зауваження 5.1. Очевидно, що нерівність Чебишева може бути записана і в іншій формі:

$$P\left\{|\xi - \mathsf{E}\xi| < \varepsilon\right\} \geq 1 - \frac{D\xi}{\varepsilon^2}.$$

Зауваження 5.2. Якщо величина ξ має абсолютний момент порядку a, то, як легко перевірити, має місце нерівність

$$P\left\{|\xi - \mathsf{E}\xi| < \varepsilon\right\} \geq 1 - \frac{\mu_a}{\varepsilon^a}. \tag{5.2}$$

Теорема 5.1. (Теорема Чебишева.) *Якщо* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *– послідовність попарно незалежних випадкових величин, що мають скінченні дисперсії, обмежені однією і тією ж сталою* C*:*

$$D\xi_1 \leq C, \ldots D\xi_2 \leq C, \ldots, D\xi_n \leq C, \ldots,$$

то, яка б не була додатна стала величина $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}\xi_k\right| < \varepsilon\right\} = 1.$$

Доведення. Із попереднього ми знаємо, що в умовах теореми

$$D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n^2}\sum_{k=1}^{n}D\xi_k,$$

звідси, згідно з умовою теореми, знаходимо, що

$$D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) \leq \frac{C}{n}.$$

Згідно з нерівністю Чебишева

$$P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n}\mathsf{E}\xi_k\right| < \varepsilon\right\} \geq 1 - \frac{D\left(\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k\right)}{\varepsilon^2} \geq 1 - \frac{C}{\varepsilon^2 n}. \tag{5.3}$$

Переходячи до границі при $n \to \infty$, дістаємо

$$\lim_{n\to\infty} P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n}\mathsf{E}\xi_k\right| < \varepsilon\right\} \geq 1.$$

а оскільки ймовірність не може бути більшою за одиницю, звідси випливає твердження теореми. □

Теорема 5.2. (Теорема Бернуллі.) *Нехай* μ– *число здійснень події* A *при* n *незалежних випробуваннях і нехай* p *є ймовірність здійснення події* A *в кожному із випробувань. Тоді яке б не було* $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1. \tag{5.4}$$

Доведення. Користуючись рівностями

$$\mathsf{E}\frac{\mu}{n} = p, \quad D\frac{\mu}{n} = \frac{pq}{n},$$

маємо

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = P\left\{\left|\frac{\mu}{n} - \mathsf{E}\frac{\mu}{n}\right| < \varepsilon\right\}.$$

Для оцінки ймовірності, яка цікавить нас, можемо скористатися нерівністю Чебишева

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \geq 1 - \frac{D\frac{\mu}{n}}{\varepsilon^2} = 1 - \frac{pq}{\varepsilon^2 n}. \tag{5.5}$$

Поклавши $n \to \infty$, знаходимо, що

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \geq 1.$$

Оскільки ймовірність взагалі не може бути більшою за одиницю, ми, тим самим, довели рівність (5.4). □

Як уже говорилося в попередньому параграфі, принципове значення теореми Бернуллі полягає в тому, що з як завгодно великою практичною впевненістю можна стверджувати, що відносна частота $\frac{\mu}{n}$ появ події A буде як завгодно мало відрізнятися від її ймовірності p, як тільки проведене число випробовувань досить велике. Ця обставина дає можливість робити висновки про величину невідомої ймовірності тієї чи іншої події, що нас цікавить, шляхом організації спеціальних випробовувань, у кожному з яких подія А може відбутися з однією й тією ж невідомою ймовірністю p. Коли ці випробовування незалежні між собою, то згідно з теоремою Бернуллі з практичною достовірністю можна вважати, що $p \approx \frac{\mu}{\mathrm{n}}$, як тільки число проведених випробовувань досить велике. Ґрунтуючись на цьому, у розд. 1 введено поняття про статистичну ймовірність і згодом не раз при розв'язанні різних прикладів його було використано. Як правило, на практиці ймовірність тієї чи іншої події доводиться визначати як статистичну ймовірність. Саме тому була проведена велика кількість експериментів для перевірки того, як теорема Бернуллі узгоджується з досвідом. Розглядалися події, ймовірності яких можна вважати з тих чи інших причин відомими і відносно яких легко провести випробовування, забезпечити їхню незалежність і сталість ймовірності в кожному з випробовувань. Усі ці випробовування дали прекрасні результати. Наведемо як приклад кілька результатів експериментів такого роду.

Приклад 5.1. Французький природознавець XVIII ст. Ж. Бюффон кинув монету 4040 разів, герб випав при цьому 2048 разів. Частота появи герба в досліді Бюффона наближено дорівнює 0,507, і лише ненабагато відрізняється від ймовірності – 0,500.

Англійський статистик К. Пірсон кинув монету 12000 разів і при цьому спостеріг 6019 випадань герба. Частота випадання герба в досліді дорівнює 0,5016.

Іншим разом він кинув монету 24000 разів і спостерігав випадання герба 12012 разів; частота випадання герба при цьому була рівною 0,5005.

Зрозуміло, що теорема Бернуллі є частковим випадком теореми Чебишева.

Ми відзначимо ще два важливі часткові випадки теореми Чебишева.

Теорема 5.3. (Теорема Пуассона.) *Якщо в послідовності незалежних випробовувань імовірність появи події A при k-му випробовуванні дорівнює p_k, то*

$$\lim_{n\to} P\left\{\left|\frac{\mu}{n} - \frac{p_1+p_2+\ldots+p_n}{n}\right| < \varepsilon\right\} = 1,$$

де, як звичайно, через μ позначено число появ події A в n перших випробовуваннях.

Доведення. Для доведення досить ввести послідовність $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ випадкових величин, що набувають значення 0 або 1 залежно від того, відбулася чи не відбулася подія A у випробовуванні з відповідним номером і зауважити, що

$$\mathsf{E}\xi_k = p_k, D\xi_k = p_k q_k \le \frac{1}{4},\ k = 1, 2, \ldots$$

$\square$

Теорема 5.4. *Якщо послідовність* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *попарно незалежних випадкових величин така, що*

а) $\mathsf{E}\xi_1 = \mathsf{E}\xi_2 = \ldots = \mathsf{E}\xi_n = \ldots = a$,

б) $D\xi_1 \leq C, D\xi_2 \leq C, \ldots, D\xi_n \leq C, \ldots$,

то, яке б не було стале число $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a\right| < \varepsilon\right\} = 1.$$

Цей частковий випадок теореми Чебишева дає обґрунтування правила середнього арифметичного, яке зазвичай вживається, коли за допомогою деяких вимірів невідомої величини хочуть мати якомога точніше уявлення про її значення. Якщо виміри не мають систематичної похибки (іншими словами, якщо математичне сподівання результатів вимірів дорівнює величині, яка вимірюється) і виміри виконані з однаковою точністю, то згідно з теоремою Чебишева, які б не були ε і η, при досить великому числі вимірювань можна стверджувати з імовірністю більшою ніж $1 - \eta$, що середнє арифметичне результатів вимірів відхиляється від вимірюваної величини не більше ніж на ε. Проте, якщо приклади допускають систематичну похибку або сам спостерігач у процесі вимірювання робить, крім випадкової, ще й систематичну похибку, то ясно, що середнє арифметичне результатів спостережень буде відрізнятися від вимірюваної величини.

Зауважимо, що доведені теореми не дають способів для обчислення ймовірності того, що відхилення частоти від імовірності або середнього від математичного сподівання не перевищить величини ε. Закон великих чисел (див. нерівності (5.5) і (5.3)) дає лише грубу оцінку цих імовірностей. Для точнішого обчислення їх треба звернутися до граничних теорем теорії ймовірностей: у випадку оцінки відхилення частоти від імовірності – до тереми Лапласа, при оцінці відхилення середнього від математичного сподівання – до теореми Ляпунова.

При деяких додаткових обмеженнях можна і безпосередньо вказати точніші оцінки вказаних імовірностей. Так, наприклад, нехай потрібно встановити скільки підкидань монети треба провести, щоб з імовірністю, не меншою 0,8, мала місце нерівність

$$\left|\frac{\mu}{n} - \frac{1}{2}\right| < 0,01.$$

Із нерівності (5.5) можна лише вивести, що 12500 випробувань для цього буде досить. Обчислення за теоремою Лапласа свідчать, що досить провести 4100 випробувань.

Теорема 5.5. (Теорема Маркова.) *Якщо послідовність випадкових величин* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *така, що при* $n \to \infty$

$$\frac{1}{n^2} D\left(\sum_{k=1}^{n} \xi_k\right) \to 0, \tag{5.6}$$

то, яке б не було додатне стале ε,

$$\lim_{n\to\infty} P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k)\right| < \varepsilon\right\} = 1.$$

Доведення. Згідно з нерівністю Чебишева

$$P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k)\right| < \varepsilon\right\} \geq 1 - \frac{D\left(\sum\limits_{k=1}^{n} \xi_k\right)}{n^2\varepsilon^2}.$$

Взявши до уваги (5.6), знаходимо, що

$$\lim_{n\to\infty} P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k)\right| < \varepsilon\right\} \geq 1.$$

Ця нерівність доводить теорему. □

Зауваження 5.3. Якщо величини $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ попарно незалежні, то умова Маркова набуває вигляду: при $n \to \infty$

$$\frac{1}{n^2}\sum_{k=1}^{n} D\xi_k \to 0.$$

Звідси видно, що теорема Чебишева є частинним випадком теореми Маркова.

Теорема 5.6. (Теорема Хінчина.) *Якщо послідовність* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *попарно незалежних однаково розподілених випадкових величин мають скінченне математичне сподівання*

$\mathsf{E}\xi_1 = \mathsf{E}\xi_2 = \ldots = \mathsf{E}\xi_n = \ldots = a$, *то для неї справджується закон великих чисел. Тобто для всякого* $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a\right| < \varepsilon\right\} = 1.$$

Ми не будемо наводити доведення цієї теореми.Лише відмітимо, що вона є наслідком теореми Колмогорова про посилений закон великих чисел.

5.2 Необхідні та достатні умови для закону великих чисел

Ми бачили, що закон великих чисел є однією з основних теорем теорії ймовірностей. Тому стає зрозумілим, чому так багато зусиль покладено на те, щоб указати найбільш широкі умови, які повинні задовольняти величини $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, щоб для них мав місце закон великих чисел.

Історія питання така. У кінці XVII і на початку XVIII ст. Яковом Бернуллі була доведена теорема, що носить його ім'я. Ця теорема була опублікована в 1713 р. у праці "Ars conjectandi" ("Майстерність припущень") лише після смерті її автора. Потім у кінці XVIII ст. С.-Д. Пуассон довів аналогічну теорему за більш загальних умов. До середини XIX ст. в цьому напрямі не було досягнуто яких-небудь успіхів. У 1866 р. російський математик П.Л.Чебишев знайшов метод (викладений у попередньому параграфі), за яким йому вдалося зовсім просто довести теорему, що включала як найпростіші частинні випадки теореми Бернуллі та Пуассона. Пізніше А.А. Марков помітив, що міркування П.Л. Чебишева дають можливість дістати більш загальні теореми. Подальші зусилля довго не давали принципових зрушень. Лише в 1926 р. А.М. Колмогоров знайшов умови, які є не лише достатніми, а й необхідними для того, щоб послідовність незалежних випадкових величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$була підпорядкована закону великих чисел. У 1928 р. О.Я. Хінчин показав, що коли випадкові величини $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ незалежні й однаково розподілені, то існування математичного сподівання $\mathsf{E}\xi_n$ є достатньою умовою для застосування закону великих чисел. В останні роки багато праць було присвячено визначенню умов, які треба накласти на залежні величини, щоб і для них мав місце закон великих чисел. Теорема Маркова належить до теорем такого роду.

Використовуючи метод Чебишева, легко зайти умову, аналогічну умові Маркова, але вже не тільки достатню, але й необхідну для застосування закону великих чисел до послідовності $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ довільних випадкових величин.

Теорема 5.7. *Для того, щоб для послідовності* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *(як завгодно залежних випадкових величин) за довільної сталої* $\varepsilon > 0$ *виконувалося співвідношення*

$$\lim_{n\to\infty} P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n} E\xi_k\right| < \varepsilon\right\} = 1, \tag{5.7}$$

необхідно і достатньо, щоб при $n \to \infty$

$$E\frac{\left(\sum\limits_{k=1}^{n}(\xi_k - E\xi_k)\right)^2}{n^2 + \left(\sum\limits_{k=1}^{n}(\xi_k - E\xi_k)\right)^2} \to 0. \tag{5.8}$$

Доведення. Припустимо спочатку, що (5.8) виконано і покажемо, що в цьому випадку виконується також (5.7). Позначимо через $\Phi_n(x)$ функцію розподілу величини

$$\eta_n = \frac{1}{n}\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k).$$

Легко перевірити такий ряд співвідношень

$$P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k)\right| \geq \varepsilon\right\} = P\{|\eta_n| \geq \varepsilon\} = \int\limits_{|x|\geq\varepsilon} d\Phi_n(x) \leq$$

$$\leq \frac{1+\varepsilon^2}{\varepsilon^2}\int\limits_{|x|\geq\varepsilon}\frac{x^2}{1+x^2}d\Phi_n(x) \leq \frac{1+\varepsilon^2}{\varepsilon^2}\int\limits_{-\infty}^{\infty}\frac{x^2}{1+x^2}d\Phi_n(x) =$$

$$= \frac{1+\varepsilon^2}{\varepsilon}\mathsf{E}\frac{\eta_n^2}{1+\eta_n^2}$$

Ця нерівність показує, що умова теореми є достатньою.

Покажемо тепер, що умова (5.8) необхідна, тобто коли має місце співвідношення (5.7), то має місце і співвідношення (5.8). Легко бачити, що

$$P\left\{|\eta_n| \geq \varepsilon\right\} = \int\limits_{|x|\geq\varepsilon} d\Phi_n\left(x\right) \geq \int\limits_{|x|\geq\varepsilon} \frac{x^2}{1+x^2} d\Phi_n\left(x\right) =$$

$$= \int\limits_{-\infty}^{\infty} \frac{x^2}{1+x^2} d\Phi_n\left(x\right) - \int\limits_{|x|<\varepsilon} \frac{x^2}{1+x^2} d\Phi_n\left(x\right) \geq$$

$$\geq \int\limits_{-\infty}^{\infty} \frac{x^2}{1+x^2} d\Phi_n\left(x\right) - \varepsilon^2 =$$

$$= \mathsf{E}\frac{\eta_n^2}{1+\eta_n^2} - \varepsilon^2 =$$

$$= \mathsf{E}\frac{\left[\sum\limits_{k=1}^{n}\left(\xi_k - \mathsf{E}\xi_k\right)\right]^2}{n^2 + \left[\sum\limits_{k=1}^{n}\left(\xi_k - \mathsf{E}\xi_k\right)\right]^2} - \varepsilon^2. \tag{5.9}$$

Припустимо, що теорема невірна і що із (5.7) не випливає(5.8). Це означає, що існує така додатна стала $a > 0$ і така нескінченна послідовність натуральних чисел n_s, що

$$\mathsf{E}\frac{\left[\sum\limits_{k=1}^{n_s}\left(\xi_k - \mathsf{E}\xi_k\right)\right]^2}{{n_s}^2 + \left[\sum\limits_{k=1}^{n_s}\left(\xi_k - \mathsf{E}\xi_k\right)\right]^2} \geq a$$

Але звідси випливає на підставі (5.9), що яке б не було $\varepsilon > 0$,

$$P\left\{|\eta_{n_s}| \geq \varepsilon\right\} \geq a - \varepsilon^2,$$

і, отже, якщо $\varepsilon < \sqrt{\frac{a}{2}}$, то

$$P\left\{|\eta_{n_s}| \geq \varepsilon\right\} \geq \frac{a}{2},\ s = 1, 2, \ldots$$

Ця нерівність суперечить співвідношенню (5.7). Отже, теорема доведена. $\square$

Зазначимо, що всі теореми, доведені в попередньому параграфі, легко випливають із щойно доведеного загального твердження.

Справді, оскільки при довільному n і довільних ξ_k має місце нерівність

$$\frac{\eta_n^2}{1+\eta_n^2} \leq \eta_n^2 = \left[\frac{1}{n}\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k)\right]^2,$$

то звідси виводимо, що з існування дисперсії у випадкових величин ξ_k випливає співвідношення

$$E\frac{\eta_n^2}{1+\eta_n^2} \leq \frac{1}{n^2} D\sum_{k=1}^{n}\xi_k.$$

Отже, якщо умова Маркова виконана, то виконана також умова (5.8), і послідовність випадкових величин ξ_1,$\xi_2, \ldots, \xi_n, \ldots$ підпорядкована закону великих чисел.

Проте слід зауважити, що в складніших випадках, коли не припускається існування скінченних дисперсій у величин ξ_k, доведена теорема для фактичної перевірки застосування закону великих чисел має досить обмежені можливості, оскільки умова (5.8) стосується не окремих доданків, а їхньої суми. Але розраховувати на те, що, не наклавши ніяких обмежень на величини ξ_k і на існуючу між ними залежність, удасться дістати умови, достатні для розв'язання задачі і в той же час зручні для застосувань, очевидно, не можна.

Якщо ж допустити, що величини ξ_1,$\xi_2, \ldots, \xi_n, \ldots$ взаємно незалежні, то можна показати, що умова (5.8) еквівалентна такій: при $n \to \infty$

$$\sum_{k=1}^{n} E\frac{\zeta_k^2}{n^2+\zeta_k^2} \to 0,$$

де позначено $\zeta_k = \xi_k - M\xi_k$.

5.2.1 Метод Монте-Карло

Закон великих чисел можна застосувати до наближеного обчислення інтегралів. В основі методу Монте-Карло лежить закон великих чисел та теорема про обчислення математичного сподівання функції від випадкової величини:

$$I = \int_0^1 f(x)dx = \mathsf{E}f(\xi),$$

де ξ – рівномірно розподілена на $[0,1]$ випадкова величина.

Теорема 5.8. (Теорема про наближення інтегралу методом Монте-Карло.) *Нехай функція f інтегровна в квадраті: $f \in L_2[0,1]$, та $\{\xi_k, k = 1, 2, \dots\}$ - послідовність незалежних рівномірно розподілених на $[0,1]$ випадкових величин. Тоді має місце збіжність*

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} f(\xi_k) - \int_0^1 f(x)dx\right| < \varepsilon\right\} = 1.$$

Доведення. Визначимо величини $\eta_k = f(\xi_k)$. Випадкові величини $\eta_k, k = 1, \dots, n$ є незалежними в сукупності, однаково розподіленими, та мають скінченні математичні сподівання

$$\mathsf{E}f(\xi_k) = \int_0^1 f(x)dx$$

і дисперсії

$$\sigma^2 = \int_0^1 f^2(x)dx - \left(\int_0^1 f(x)dx\right)^2.$$

Тому для них виконується закон великих чисел. Більше того, можна оцінити швидкість збіжності в середньому квадратичному

$$\mathsf{E}\left(\frac{1}{n}\sum_{k=1}^{n} f(\xi_k) - I\right)^2 = D\left(\frac{1}{n}\sum_{k=1}^{n} f(\xi_k)\right) = \frac{\sigma^2}{n} = O\left(\frac{1}{n}\right).$$

На відміну від схем наближеного обчислення інтегралів із регулярно розміщеними вузлами метод Монте-Карло є збіжним і для розривних функцій. □

Таким чином теореми про закон великих чисел дають теоретичне обгрунтування наступного алгоритму обчислення інтеграла

$$I = \int_0^1 f(x)dx.$$

1. Моделюємо послідовність $\{\xi_k, k = 1, 2, \dots, n,\}$ незалежних рівномірно розподілених на $[0,1]$ випадкових величин.

2. Обчислюємо наближене значення інтеграла

$$\int_0^1 f(x)dx \simeq \frac{1}{n}\sum_{k=1}^{n} f(\xi_k).$$

Метод Монте-Карло особливо ефективний при обчисленні інтегралів великої кратності.

5.3 Збіжність за ймовірністю та її властивості

Означення 5.2. Послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається за ймовірністю до величини ξ (позначення: $\xi = P\cdot\lim_{n\to\infty}\xi_n$, а також $\xi_n \xrightarrow{P} \xi, n \to \infty$), якщо для кожного $\varepsilon > 0$

$$P\{|\xi_n - \xi| \geq \varepsilon\} \to 0 \quad \text{при } n \to \infty.$$

Зауваження 5.4. Якщо одночасно $\xi_n \xrightarrow{P} \xi$ та $\xi_n \xrightarrow{P} \eta$, то $P\{|\xi - \eta| \geq \varepsilon\} = 0$ для всіх $\varepsilon > 0$, звідки $P\{\xi - \eta = 0\} = 1$, тобто $\xi - \eta = 0$ м.н., тобто границя за ймовірністю визначена однозначно з точністю до рівності майже напевне.

Лема 5.2. *Нехай $\xi_n \xrightarrow{P} \xi, n \to \infty$, та $f\colon \mathbb{R} \to \mathbb{R}$ – неперервна функція. Тоді*

$$f(\xi_n) \xrightarrow{P} f(\xi),\ n \to \infty.$$

Доведення. Для довільного $\varepsilon > 0$ та $c > 0$ оберемо $\delta > 0$ так, щоб $|f(x_1) - f(x_2)| \leq \varepsilon$ при $|x_1 - x_2| \leq \delta$, $|x_1| \leq c$. (Таке $\delta > 0$ можна вибрати оскільки функція f рівномірно неперервна на обмеженій множині.) Тоді

$$P\{|f(\xi_1) - f(\xi_2)| > \varepsilon\} \leq P\{|\xi| > c\} + P\{|\xi_1 - \xi_2| > \delta\}.$$

Отже

$$\overline{\lim}_{n\to\infty} P\{|f(\xi_1) - f(\xi_2)| > \varepsilon\} \leq P\{|\xi| > c\}.$$

Оскільки величина в правій частині вибором $c > 0$ може бути як завгодно малою, то

$$\overline{\lim}_{n\to\infty} P\{|f(\xi_1) - f(\xi_2)| > \varepsilon\} = 0.$$

□

Лема 5.3. *Нехай* $\xi_n \xrightarrow{P} \xi, n \to \infty$, *та* $P\{|\xi_n| \le c\} = 1$ *Тоді*

$$\lim_{n\to\infty} E\xi_n = E\xi.$$

Доведення. Зауважимо що $P\{|\xi| \le c\} = 1$. Дійсно, визначимо неперервну функцію $f(x)$ так щоб $f(x) = 0$ при $|x| \le c$ та $f(x) > 0$ при $|x| > c$. Тоді $P\{f(\xi_n) = 0\} = 1$ та $f(\xi_n) \xrightarrow{P} f(\xi),\ n \to \infty$. Отже $P\{f(\xi_n) = 0\} = 1$ та $P\{|\xi| \le c\} = 1$.

Визначимо випадкові величини $\chi_n(\delta) = 1$, якщо $|\xi_n - \xi| > \delta$ та $\chi_n(\delta) = 0$, якщо $|\xi_n - \xi| \le \delta$. Тоді

$$|\xi_n - \xi| \le \delta + 2c\chi_n(\delta).$$

Таким чином

$$|\mathsf{E}\xi_n - \mathsf{E}\xi| \le \mathsf{E}|\xi_n - \xi| \le \delta + 2c\mathsf{E}\chi_n(\delta) \le \delta + 2cP\{|\xi_n - \xi| > \delta\}.$$

Тому

$$\lim_{n\to\infty} |\mathsf{E}\xi_n - \mathsf{E}\xi| \le \delta$$

для всіх $\delta > 0$. □

Зауваження 5.5. Нехай $\xi_n \xrightarrow{P} \xi, n \to \infty$. Тоді $|\xi_n - \xi| \xrightarrow{P} 0, n \to \infty$. Якщо $P\{|\xi_n - \xi| \le c\} = 1$ при деякому $c > 0$ то

$$\lim_{n\to\infty} \mathsf{E}|\xi_n - \xi| = 0.$$

Означення 5.3. Послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ в середньому, якщо для всіх n існують моменти $\mathsf{E}\xi_n < \infty$, $\mathsf{E}\xi < \infty$ та

$$\lim_{n\to\infty} \mathsf{E}|\xi_n - \xi| = 0.$$

У цьому випадку

$$\lim_{n\to\infty} \mathsf{E}\xi_n = \mathsf{E}\xi.$$

Лема 5.4. *Нехай послідовність випадкових величин* $\{\xi_n, n = 1, 2, \dots\}$ *збігається до випадкової величини* ξ *в середньому. Тоді послідовність* $\{\xi_n, n = 1, 2, \dots\}$ *збігається до величини* ξ *за ймовірністю:* $\xi_n \xrightarrow{P} \xi, n \to \infty$.

Доведення. Ця властивість є наслідком нерівності Чебишева

$$P\{|\xi_n - \xi| > \varepsilon\} \leq \frac{1}{\varepsilon}\mathsf{E}|\xi_n - \xi|.$$

□

Означення 5.4. Послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ в середньоквадратичному, якщо для всіх n існують моменти $\mathsf{E}\xi_n^2 < \infty$, $\mathsf{E}\xi^2 < \infty$ та

$$\lim_{n\to\infty} \mathsf{E}|\xi_n - \xi|^2 = 0.$$

Лема 5.5. *Нехай послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ в середньоквадратичному. Тоді послідовність $\{\xi_n, n = 1, 2, \dots\}$ збігається до величини ξ в середньому, а тож за ймовірністю $\xi_n \xrightarrow{P} \xi, n \to \infty$.*

Доведення. Ця властивість є наслідком нерівності

$$\mathsf{E}|\xi_n - \xi| \leq \left(\mathsf{E}|\xi_n - \xi|^2\right)^{\frac{1}{2}}.$$

□

Лема 5.6. *Нехай послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ в середньоквадратичному. Тоді*

$$\mathsf{E}\xi_n \to \mathsf{E}\xi, \ \text{та} \ \ \mathsf{E}\xi_n^2 \to \mathsf{E}\xi^2.$$

Доведення. Перша властивість є наслідком поперенього твердження.

Для доведення другої зауважимо, що

$$\xi_n^2 = [\xi + (\xi_n - \xi)]^2 \leq 2[\xi^2 + (\xi_n - \xi)^2].$$

Тому

$$\mathsf{E}\xi_n^2 \leq 2[\mathsf{E}\xi^2 + \mathsf{E}(\xi_n - \xi)^2].$$

Отже

$$\begin{gathered}|\mathsf{E}\xi_n^2 - \mathsf{E}\xi^2| = |\mathsf{E}(\xi_n - \xi)(\xi_n + \xi)| \leq \\ \leq [\mathsf{E}(\xi_n - \xi)^2]^{\frac{1}{2}}[\mathsf{E}(\xi_n + \xi)^2]^{\frac{1}{2}} \leq \\ \leq [\mathsf{E}(\xi_n - \xi)^2]^{\frac{1}{2}}[2\mathsf{E}(\xi_n^2 + \xi^2)]^{\frac{1}{2}} \leq \\ \leq [\mathsf{E}(\xi_n - \xi)^2]^{\frac{1}{2}}[6\mathsf{E}\xi^2 + 4\mathsf{E}(\xi_n - \xi)^2)]^{\frac{1}{2}}\end{gathered}$$

Останній вираз прямує до нуля. □

Лема 5.7. *Для того щоб послідовність випадкових величин* $\{\xi_n, n = 1, 2, \dots\}$ *збігалась в середньоквадратичному до сталої величини* c *необхідно і достатньо, щоб*

$$\mathsf{E}\xi_n \to c, \text{ та } D\xi_n \to 0.$$

Доведення. Якщо послідовність випадкових величин $\{\xi_n\}$ збігається до сталої величини c в середньоквадратичному, то

$$\mathsf{E}\xi_n \to \mathsf{E}c = c, \ \mathsf{E}\xi_n^2 \to c^2.$$

Тому

$$D\xi_n = \mathsf{E}\xi_n^2 - (\mathsf{E}\xi_n)^2 \to 0.$$

Навпаки. Якщо

$$\mathsf{E}\xi_n \to c, \text{ та } D\xi_n \to 0,$$

то

$$\mathsf{E}(\xi_n - c)^2 = \mathsf{E}(\xi_n - \mathsf{E}\xi_n)^2 + (\mathsf{E}\xi_n - c)^2 \to 0.$$

□

5.4 Збіжність з ймовірністю 1

Означення 5.5. Послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ з імовірністю 1, або ж збігається майже напевне, позначення $\xi_n \xrightarrow{P1} \xi$, $P\{\xi_n \to \xi\} = 1$, $P\{\lim \xi_n = \xi\} = 1$, якщо

$$P\{\omega \colon \exists \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\} = 1.$$

За означенням, послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ майже напевне, якщо подія

$$A = \{\exists \lim \xi_n = \xi\} = \cap_{k\geq 1} \cup_{m\geq 1} \cap_{n\geq m} \{|\xi_n - \xi| \leq 1/k\}$$

має одиничну ймовірність: $P(A) = 1$.

Позначимо через

$$A_{mk} = \cap_{n\geq m} \{|\xi_n - \xi| \leq 1/k\}$$

подію, яка полягає у тому, що нерівність

$$\{|\xi_n - \xi| \leq 1/k\}$$

виконується для всіх $n \geq m$.

Через $B_k = \cup_{m\geq 1} A_{mk}$ позначимо подію, яка полягає у тому, що існує таке m, що вказана нерівність справджується для всіх $n \geq m$.

Через $A = \cap_{k\geq 1} B_k$ позначимо подію, яка полягає у тому, що для всіх k існує таке m, що вказана нерівність справджується для всіх $n \geq m$.

Якщо $P\{A\} = P\{\cap_{k\geq 1} B_k\} = 1$, то $P\{B_k\} = 1$ для всіх k. І навпаки, якщо $P\{B_k\} = 1$ для всіх k, то $P\{A\} = P\{\cap_{k\geq 1} B_k\} = 1$. Таким чином, для того щоб $\xi_n \xrightarrow{P1} \xi$ необхідно і достатньо, щоб для всіх k

$$P\{B_k\} = P\{\cup_{m\geq 1} \cap_{n\geq m} \{|\xi_n - \xi| \leq 1/k\}\} = 1. \tag{5.10}$$

Оскільки послідовність

$$A_{mk} = \cap_{n\geq m} \{|\xi_n - \xi| \leq 1/k\}$$

зростаюча як функція від m, то (5.10) виконується тоді і тільки тоді, коли для всіх k

$$\lim_{m\to\infty} P\{\cap_{n\geq m} \{|\xi_n - \xi| \leq 1/k\}\} = 1.$$

Оскільки число $1/k$ можна замінити на довільне $\varepsilon > 0$, то ми переконуємося у справедливості такого твердження.

Лема 5.8. *Для того щоб послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігалась до випадкової величини ξ з імовірністю 1, необхідно і достатньо, щоб для всіх $\varepsilon > 0$ виконувалась умова*

$$\lim_{m\to\infty} P\{\cap_{n\geq m} \{|\xi_n - \xi| \leq \varepsilon\}\} = 1.$$

Зауваження 5.6. Оскільки

$$P\{|\xi_n - \xi| \leq \varepsilon\} \geq P\{\cap_{n\geq m} \{|\xi_n - \xi| \leq \varepsilon\}\},$$

то із збіжності послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ до випадкової величини ξ з імовірністю 1 виливає, що послідовність $\{\xi_n, n = 1, 2, \dots\}$ збігається до випадкової величини ξ за імовірністю. У цьому випадку

$$P\{|\xi_n - \xi| \leq \varepsilon\} \to 1.$$

Наслідок 5.1. *Умова леми еквівалентна такій умові: для будь-якого $\varepsilon > 0$*

$$\lim_{m\to\infty} P\{\cup_{n\geq m}\{|\xi_n - \xi| > \varepsilon\}\} = 0.$$

Наслідок 5.2. *Для того щоб послідовність випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ збігалась до випадкової величини ξ з імовірністю 1, достатньо, щоб для всіх $\varepsilon > 0$ виконувалась умова*

$$\sum_{n=1}^{\infty} P\{|\xi_n - \xi| > \varepsilon\} < \infty.$$

Лема 5.9. (Теорема Бореля.) *Нехай μ позначає кількість успіхів у послідовності n незалежних дослідів Бернуллі у кожному з яких ймовірність успіху дорівнює p. Тоді при $n \to \infty$*

$$P\left\{\frac{\mu}{n} \to p\right\} = 1.$$

Доведення. Щоб довести твердження досить довести збіжність ряду

$$\sum_{n=1}^{\infty} P\left\{\left|\frac{\mu}{n} - p\right| > \frac{1}{r}\right\} < \infty$$

для кожного натурального числа r. Скористаємося нерівностю Чебишева

$$P\left\{\left|\frac{\mu}{n} - p\right| > \frac{1}{r}\right\} \leq r^4\, \mathsf{E}\left(\frac{\mu}{n} - p\right)^4$$

щоб оцінити кожний доданок у сумі ряду. Позначимо через ξ_k кількість успіхів в k-му досліді Бернуллі. Тоді

$$\frac{\mu}{n} - p = \frac{1}{n}\sum_{k=1}^{n}(\xi_k - p).$$

Обчислимо

$$\mathsf{E}\left(\frac{\mu}{n} - p\right)^4 = \frac{1}{n^4}\sum_{k=1}^{n}\sum_{j=1}^{n}\sum_{i=1}^{n}\sum_{l=1}^{n}\mathsf{E}(\xi_k - p)(\xi_j - p)(\xi_i - p)(\xi_l - p).$$

Оскільки $\mathsf{E}(\xi_k - p) = 0$, то у цій суммі відмінні від нуля лише доданки, які мають вигляд $\mathsf{E}(\xi_k - p)^4$ та $\mathsf{E}(\xi_k - p)^2(\xi_j - p)^2$. При цьому

$$\mathsf{E}(\xi_k - p)^4 = pq(p^3 + q^3),$$

$$\mathsf{E}\,(\xi_k - p)^2\,(\xi_j - p)^2 = p^2q^2,\; k \neq j.$$

Число доданків першого виду дорівнює n. Число доданків другого виду дорівнює $3n(n-1)$. Таким чином

$$\mathsf{E}\left(\frac{\mu}{n} - p\right)^4 = \frac{pq}{n^4}[n(p^3+q^3) + 3pq(n^2-n)] \leq \frac{1}{4n^2}.$$

Звідси випливає збіжність ряду і справедливість твердження леми. $\square$

Цю теорему, яку Е. Борель довів у 1909 р., можна сформулювати у такому вигляді

$$P\left\{\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}\xi_k \to 0\right\} = 1.$$

Тобто із ймовірністю 1 різниця між середнім арифметичним незалежних випадкових величин, які приймають значення 1 та 0 з ймовірностями p та $q = 1 - p$, та середнім арифметичним математисних сподівань цих випадкових величин пряму до нуля при зростанні $n \to \infty$.

5.5 Посилений закон великих чисел

Твердження теореми Бореля ми отримаємо пізніше як частинний випадок теореми Колмогорова, яка вказує на умови при яких для послідовності незалежних випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ справджується посилений закон великих чисел.

Означення 5.6. Для послідовності випадкових величин $\{\xi_n, n = 1, 2, \dots\}$ виконується посилений закон великих чисел, якщо різниця між середнім арифметичним випадкових величин та середнім арифметичним математисних сподівань цих випадкових величин задовольняє граничне співвідношення при зростанні $n \to \infty$

$$P\left\{\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}\xi_k \to 0\right\} = 1.$$

Як і для звичайного закону великих чисел, у випадку однаково розподілених випадкових величин це співвідношення еквівалентне збіжності середніх арифметичних випадкових величин до математичного сподівання із ймовірністю 1, тобто

$$P\left\{\frac{1}{n}\sum_{k=1}^{n}\xi_k \to \mathsf{E}\xi_1\right\} = 1.$$

Зараз ми встановимо одну нерівність Колмогорова, яка є уточненням нерівності Чебишева.

Теорема 5.9. (Нерівність Колмогорова.) *Якщо*

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

– взаємно незалежні випадкові величини, що мають скінченні дисперсії, то ймовірність сумісного здійснення нерівностей

$$\left|\sum_{s=1}^{k}(\xi_s - \mathsf{E}\xi_s)\right| < \varepsilon, \quad k = 1, 2, \ldots, n$$

не менша ніж

$$1 - \frac{1}{\varepsilon^2}\sum_{k=1}^{n} D\xi_k.$$

Тобто

$$P\left\{\max_{1\le k\le n} |S_k| \ge \varepsilon\right\} \le \frac{1}{\varepsilon^2} DS_n.$$

Доведення. Введемо такі позначення:

$$\eta_j = \xi_j - \mathsf{E}\xi_j, \quad S_k = \sum_{j=1}^{k}\eta_j.$$

Позначимо далі буквою E_k подію, що полягає в сумісному здійсненні таких нерівностей:

$$|S_j| < \varepsilon, \ \ j \le k-1, \quad |S_k| \ge \varepsilon.$$

E_0 означає подію, яка полягає в тому, що $|S_j| < \varepsilon$ для $1 \le j \le n$. Очевидно, що подія, яка полягає в тому, що хоча б при одному k буде виконана нерівність

$$|S_k| \ge \varepsilon, \quad k = 1, 2, \ldots, n, \tag{5.11}$$

дорівнює $\cup_{k=1}^{n} E_k$ і, отже, в силу несумісності подій E_k

$$P\left\{\max_{1\leq k\leq n} |S_k| \geq \varepsilon\right\} = \sum_{k=1}^{n} P\{E_k\}.$$

Можемо записати

$$\mathrm{D}S_n = \sum_{k=0}^{n} P\{E_k\}\,\mathrm{D}\left(S_n|E_k\right) \geq \sum_{k=1}^{n} P\{E_k\}\,\mathrm{D}\left(S_n|E_k\right).$$

Оскільки очевидно, що

$$\mathrm{D}\left(S_n|E_k\right) = \mathsf{E}\left[\left(S_k^2 + 2(S_n - S_k)S_k + (S_n - S_k)^2\right)|E_k\right] \geq$$

$$\geq \mathsf{E}\left[\left(S_k^2 + 2(S_n - S_k)S_k + 0\right)|E_k\right] = \mathsf{E}\left[\left(S_k^2\right)|E_k\right]$$

і в силу незалежності величин η_j мають місце рівності

$$\mathsf{E}\left[(S_n - S_k)S_k|E_k\right] = 0$$

а згідно з (5.11) – нерівність

$$\mathsf{E}\left[S_k^2|E_k\right] \geq \varepsilon^2, \quad k \geq 1,$$

то ми можемо записати, що

$$\mathrm{D}S_n \geq \varepsilon^2 \sum_{k=1}^{n} P\{E_k\}.$$

Звідси

$$\sum_{k=1}^{n} P\{E_k\} = \mathrm{P}\left\{\max_{1\leq \mathrm{k} \leq \mathrm{n}} |S_k| \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2}\mathrm{D}S_n.$$

Нерівність Колмогорова доведена. $\square$

Теорема 5.10. (Теорема Колмогорова.) *Якщо послідовність взаємно незалежних випадкових величин*

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

задовольняє умову

$$\sum_{n=1}^{\infty} \frac{D\xi_n}{n^2} < +\infty,$$

то вона підпорядкована посиленому закону великих чисел. Тобто

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}(\xi_k - \mathsf{E}\xi_k) = 0\right\} = 1.$$

Доведення. Будемо вважати, що $\mathsf{E}\xi_k = 0$. Покладемо

$$\eta_n = \max_{m\le 2^n}\left|\sum_{k=1}^{m}\xi_k\right|.$$

Оскільки

$$\left|\frac{1}{m}\sum_{k=1}^{m}\xi_k\right| \le 2^{-n+1}\max_{m\le 2^n}\left|\sum_{k=1}^{m}\xi_k\right| = 2^{-n+1}\eta_n$$

при

$$2^{n-1} \le m \le 2^n,$$

то для доведення теореми досить показати, що

$$P\left\{\lim_{n\to\infty}2^{-n}\eta_n = 0\right\} = 1.$$

Це співвідношення виконується якщо для всякого $\varepsilon > 0$

$$\sum_{k=1}^{\infty}P\left\{2^{-n}\eta_n > \varepsilon\right\} < \infty.$$

Що довести збіжність цього ряду скористаємося нерівністю Колмогорова. Маємо

$$P\left\{2^{-n}\eta_n > \varepsilon\right\} = P\left\{\eta_n > 2^n\varepsilon\right\} \le \frac{1}{2^{2n}\varepsilon^2}\sum_{k=1}^{2^n}D\xi_k.$$

Тому

$$\sum_{n=1}^{\infty}P\left\{2^{-n}\eta_n > \varepsilon\right\} \le \varepsilon^{-2}\sum_{n=1}^{\infty}2^{-2n}\sum_{k=1}^{2^n}D\xi_k \le$$

$$\le \varepsilon^{-2}\sum_{k=1}^{\infty}D\xi_k\sum_{2^n>k}4^{-n} \le \varepsilon^{-2}\sum_{k=1}^{\infty}D\xi_k\frac{4}{3k^2} < \infty.$$

Теорема доведена. $\square$

Зауваження 5.7. Зауважимо, що коли дисперсії випадкових величин ξ_n обмежені однією й тією ж сталою C в сукупності, то послідовність взаємно незалежних величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ підпорядкована посиленому закону великих чисел.

Для однаково розподілених незалежних випадкових величин існує критерій Колмогорова виконання посиленого закону великих чисел.

Теорема 5.11. (Теорема Колмогорова.) *Нехай $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ послідовність незалежних однаково розподілених випадкових величин що мають скінченні математичні сподівання $\mathsf{E}\xi_k = a, k = 1, 2, \ldots$. Тоді вона підпорядкована посиленому закону великих чисел. Тобто*

$$P\left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \xi_k = a \right\} = 1.$$

Доведення. Визначимо випадкові величини ξ'_n за правилом $\xi'_n = \xi_n$ при $|\xi_n| \le n$ та $\xi'_n = 0$ при $|\xi_n| > n$. Нехай $\xi''_n = \xi_n - \xi'_n$. За умов теореми інтеграл

$$\int |x| dF(x) < \infty,$$

де $F(x)$- функція розподілу випадкової величини ξ_1 скінченний. Тому

$$\sum_{n=1}^{\infty} P\{|\xi_1| > n\} = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P\{k < |\xi_1| \le k+1\} =$$

$$= \sum_{n=1}^{\infty} kP\{k < |\xi_1| \le k+1\} \le$$

$$\le \sum_{n=1}^{\infty} \int_{k<x\le k+1} |x| dF(x) < \int |x| dF(x) < \infty. \qquad (5.12)$$

Далі оцінимо дисперсії величин ξ'_n

$$D\xi'_n \le \mathsf{E}(\xi'_n)^2 = \int_{-n}^{n} x^2 dF(x) \le \sum_{k=0}^{n} (k+1)^2 P\{k < |\xi_1| \le k+1\}$$

а також

$$\sum_{n=1}^{\infty} \frac{D\xi'_n}{n^2} \le \sum_{n=1}^{\infty} \sum_{k=0}^{n} \frac{(k+1)^2}{n^2} P\{k < |\xi_1| \le k+1\} \le$$

$$\sum_{k=0}^{\infty} P\{k < |\xi_1| \le k+1\}(k+1)^2 \sum_{n=k}^{\infty} \frac{1}{n^2}.$$

Оскільки

$$\sum_{n=k}^{\infty} \frac{1}{n^2} < \frac{1}{k^2} + \frac{1}{k} < \frac{2}{k},$$

то в силу

$$\sum_{n=1}^{\infty} \frac{D\xi_n'}{n^2} < \infty.$$

Отже послідовність ξ_n' задовольняє умови посиленого закону великих чисел. Тобто

$$P\left\{\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\xi_k' - \mathsf{E}\xi_k') = 0\right\} = 1.$$

За теоремою Штольца

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathsf{E}\xi_k' = \lim_{n\to\infty} \mathsf{E}\xi_n' = \lim_{n\to\infty} \int_{-n}^{n} x dF(x) = a.$$

Отже

$$P\left\{\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \xi_k' = a\right\} = 1.$$

Нам залишається показати, що ймовірність принаймі однієї нерівності $\xi_n \ne \xi_n'$ для $n \ge N$ прямує до нуля при $N \to \infty$. Дійсно,

$$P\{\xi_n \ne \xi_n' \text{ при деякому } n \ge N\} \le \sum_{n\ge N} P\{\xi_n \ne \xi_n'\} =$$

$$= \sum_{n\ge N} P\{|\xi_n| > n\} \le \sum_{n=N}^{\infty} (n-N+1) P\{n < |\xi_n| \le n+1\} \le$$

$$\le \sum_{n=N}^{\infty} n \int_{n<|x|\le n+1} dF(x) \le \sum_{n=N}^{\infty} \int_{n<|x|\le n+1} |x| dF(x)$$

$$= \int_{|x|>N} |x| dF(x).$$

Величина такого інтеграла може бути зроблена як завгодно малою якщо N вибирати достатньо великим. $\square$

6 Характеристичні функції

У попередніх розділах ми бачили, що в теорії ймовірностей широко використовуються апарат математичного аналізу. Просте розв'язання багатьох задач теорії ймовірностей, особливо пов'язаних з підсумовуванням незалежних випадкових величин, удається дістати з допомогою характеристичних функцій, теорія яких широко розвинена в аналізі й відома під назвою "Перетворення Фур'є". Цей розділ присвячено викладу основних властивостей характеристичних функцій.

6.1 Означення та найпростіші властивості характеристичних функцій

Означення 6.1. Характеристичною функцією $f(t)$ випадкової величини ξ називається математичне сподівання випадкової величини $e^{it\xi}$, де t – дійсна змінна:

$$f(t) = \mathsf{E}e^{it\xi} \tag{6.1}$$

Якщо $F(x)$ є функцією розподілу величини ξ, то характеристична функція обчислюється за формулою

$$f(t) = \int\limits_{-\infty}^{\infty} e^{itx} dF(x). \tag{6.2}$$

Із того, що $\left|e^{itx}\right| = 1$ при всіх дійсних t, випливає існування інтеграла (6.2) для всіх функцій розподілу. Таким чином, характеристична функція може бути визначена для кожної випадкової величини.

Теорема 6.1. *Характеристична функція рівномірно неперервна на всій прямій і задовольняє такі співвідношення:*

$$f(0) = 1; \quad |f(t)| \leq 1, \quad (-\infty < t < +\infty). \tag{6.3}$$

Доведення. Співвідношення (6.3) легко випливають з означення характеристичної функції. Справді, із (6.2) знаходимо

$$f(0) = \int\limits_{-\infty}^{\infty} 1 \cdot dF(x) = 1,$$

а

$$|f(t)| = \left| \int\limits_{-\infty}^{\infty} e^{itx} dF(x) \right| \le \int\limits_{-\infty}^{\infty} \left| e^{itx} \right| dF(x) = \int\limits_{-\infty}^{\infty} dF(x) = 1.$$

Залишається довести рівномірну неперервність функції $f(t)$. Для цього розглянемо різницю:

$$f(t+h) - f(t) = \int\limits_{-\infty}^{\infty} e^{itx} \left(e^{ihx} - 1 \right) dF(x)$$

й оцінимо її модуль. Маємо

$$|f(t+h) - f(t)| \le \int\limits_{-\infty}^{\infty} \left| e^{ihx} - 1 \right| dF(x).$$

Нехай $\varepsilon > 0$ – довільне; виберемо настільки велике A, щоб

$$\int\limits_{|x|>A} dF(x) < \frac{\varepsilon}{4},$$

та підберемо таке мале h, щоб для $|x| \le A$

$$\left| e^{ihx} - 1 \right| < \frac{\varepsilon}{2}.$$

Тоді

$$|f(t+h) - f(t)| \le \int\limits_{-A}^{A} \left| e^{ihx} - 1 \right| dF(x) + 2 \int\limits_{|x| \ge A} dF(x) < \varepsilon.$$

$\square$

Теорема 6.2. *Якщо $\eta = a\xi + b$, де a і b – сталі, то*

$$f_\eta(t) = f_\xi(at) e^{ibt}.$$

Теорема 6.3. *Характеристична функція суми двох незалежних випадкових величин дорівнює добуткові їхніх характеристичних функцій.*

Доведення. Нехай ξ та η – незалежні випадкові величини і $\zeta = \xi + \eta$. Тоді, очевидно, разом із ξ та η незалежні також випадкові величини $e^{it\xi}$ та $e^{it\eta}$. Звідси випливає, що

$$\mathsf{E}e^{it\zeta} = \mathsf{E}e^{it(\xi+\eta)} = \mathsf{E}e^{it\xi} \cdot \mathsf{E}e^{it\eta}.$$

Ця рівність доводить теорему. $\square$

Теорема 6.4. *Якщо $\zeta = \xi_1 + \xi_2 + \ldots + \xi_n$, причому кожний доданок ξ_k не залежить від суми попередніх доданків $\xi_1 + \xi_2 + \ldots + \xi_{k-1}$, то характеристична функція величини ζ дорівнює добуткові характеристичних функцій доданків.*

Доведення. Позначивши через η_n суму $\xi_1 + \xi_2 + \ldots + \xi_{n-1}$, можемо застосувати попередню теорему. Тому

$$f_\zeta(t) = f_{\eta_n}(t) \cdot f_{\xi_n}(t),$$

де $f_\zeta(t)$, $f_{\eta_n}(t)$ і $f_{\xi_n}(t)$ означають характеристичні функції випадкових величин ζ, η_n та ξ_n. Із цієї ж теореми випливає, що

$$f_{\eta_n}(t) = f_{\eta_{n-1}}(t) \cdot f_{\xi_{n-1}}(t)$$

, де $\eta_{n-1} = \xi_1 + \xi_2 + \ldots + \xi_{n-2}$.

Продовжуючи ці міркування, приходимо до доведення нашого твердження. $\square$

Теорема 6.5. *Якщо перший абсолютний момент випадкової величини ξ скінченний, то характеристична функція $f(t)$ величини ξ диференційовна і має місце рівність*

$$f(t) = 1 + it\mathsf{E}\xi + o(t),\ t \to).$$

Доведення. З нерівності

$$|e^{itx} - 1 - itx| \le 2|tx|,\ \forall tx \in \mathbb{R}$$

випливає, що до інтегралу

$$(f(t) - 1 - it\mathsf{E}\xi)t^{-1} = \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)t^{-1} dF(x)$$

можна застосувати теорему Лебега про мажоровану збіжність при $t \to 0$ тому що

$$\int_{-\infty}^{\infty} 2|x| dF(x) = 2\mathsf{E}\xi < \infty.$$

Оскільки підінтегральна функція в правій частині поточково прямує до нуля, то права частина прямує до нуля при $t \to 0$, звідки дістанемо співвідношення теореми. $\square$

Теорема 6.6. *Якщо другий абсолютний момент випадкової величини ξ скінченний, то характеристична функція $f(t)$ величини ξ диференційована 2 рази і має місце рівність*

$$f(t) = 1 + itE\xi - \frac{t^2}{2}E\xi^2 + o(t^2),\ t \to 0.$$

Доведення. З нерівності

$$\left| e^{itx} - 1 - itx + \frac{t^2x^2}{2} \right| \le |tx|^2,\ \forall tx \in \mathbb{R}$$

випливає, що до інтегралу

$$\left(f(t) - 1 - it\mathsf{E}\xi + \frac{t^2}{2}\mathsf{E}\xi^2 \right) t^{-2} =$$

$$= \int_{-\infty}^{\infty} \left(e^{itx} - 1 - itx + \frac{t^2x^2}{2} \right) t^{-2} dF(x)$$

можна застосувати теорему Лебега про мажоровану збіжність при $t \to 0$ тому що

$$\int_{-\infty}^{\infty} |x|^2 dF(x) = \mathsf{E}|\xi|^2 < \infty.$$

Оскільки підінтегральна функція в правій частині поточково прямує до нуля, то права частина прямує до нуля при $t \to 0$, звідки дістанемо співвідношення теореми. $\square$

Застосування характеристичних функцій в значній мірі базується на властивості, сформульованої в теоремі 6.3. Додавання незалежних випадкових величин призводить до дуже складної операції – композиції функцій розподілу доданків. Для характеристичних функцій ця складна операція заміняється звичайним множенням характеристичних функцій доданків.

Теорема 6.7. *Якщо абсолютний момент n-го порядку випадкової величини ξ скінченний, то характеристична функція $f(t)$ величини ξ диференційовна n разів, причому при $k \le n$ має місце рівність*

$$f^{(k)}(0) = i^k E\xi^k \tag{6.4}$$

Доведення. Формальне k-кратне $(k \le n)$ диференціювання характеристичної функції призводить до формальної рівності

$$f^{(k)}(t) = i^k \int\limits_{-\infty}^{\infty} x^k e^{itx} dF(x) \tag{6.5}$$

Але за умовою теореми

$$\left| \int\limits_{-\infty}^{\infty} x^k e^{itx} dF(x) \right| \le \int\limits_{-\infty}^{\infty} |x|^k dF(x) \le \int\limits_{-1}^{1} |x|^k dF(x) + \int\limits_{|x| \ge 1} |x|^k dF(x) \le$$

$$\le \int\limits_{-1}^{1} |x|^k dF(x) + \int\limits_{-\infty}^{\infty} |x|^n dF(x) < +\infty.$$

Звідси випливає рівномірна збіжність інтеграла (6.5) і законність диференціювання. Поклавши в (6.5) $t = 0$, знаходимо, що

$$f^{(k)}(0) = i^k \int\limits_{-\infty}^{\infty} x^k dF(x).$$

□

Ця рівність дозволяє дуже просто визначити моменти випадкової величини через характеристичну функцію.

Математичне сподівання і дисперсія дуже просто виражаються за допомогою похідних від логарифма характеристичної функції. Справді, покладемо

$$\psi(t) = \ln f(t),$$

тоді

$$\psi'(t) = \frac{f'(t)}{f(t)}$$

і

$$\psi''(t) = \frac{f''(t) \cdot f(t) - [f'(t)]^2}{f^2(t)}.$$

Беручи до уваги $f(0) = 1$ і рівність (6.4), знаходимо

$$\psi'(0) = f'(0) = i\mathsf{E}\xi,$$

$$\psi''(0) = f''(0) - [f'(0)]^2 = i^2 \mathsf{E}\xi^2 - [i\mathsf{E}\xi]^2 = -D\xi.$$

Звідси

$$\mathsf{E}\xi = \frac{1}{i}\psi'(0)\,, D\xi = -\psi''(0) \tag{6.6}$$

Означення 6.2. Похідні $\psi^{(k)}(0)$ логарифма характеристичної функції в точці 0, поділені на i^k, називаються **семіінваріантами** відповідної випадкової величини.

При додаванні незалежних випадкових величин їхні семіінваріанти також додаються: справді, нехай ξ та η – незалежні випадкові величини і $\zeta = \xi + \eta$. Оскільки $f_\zeta(t) = f_\xi(t) \cdot f_\eta(t)$, то $\psi_\zeta(t) = \psi_\xi(t) + \psi_\eta(t)$. Звідси шляхом k-кратного диференціювання в точці 0 та ділення на i^k можна вивести справедливість нашого твердження. Першими двома семіінваріантами є середнє значення та дисперсія, тобто момент першого порядку і певна раціональна функція від моментів першого і другого порядків. Шляхом обчислення можна переконатися в тому, що і кожний k-й семіінваріант є раціональна функція від моментів порядку $\le k$.

Розглянемо кілька прикладів обчислення характеристичних функцій.

Приклад 6.1. Випадкова величина ξ розподілена за нормальним законом з математичним сподіванням a та дисперсією σ^2. Характеристична функція величини ξ (за означенням) обчислюється за формулою

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

Заміною

$$z = \frac{x-a}{\sigma} - it\sigma$$

зводимо $f(t)$ до вигляду

$$f(t) = e^{-\frac{t^2\sigma^2}{2} + iat} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty - it\sigma}^{\infty - it\sigma} e^{-\frac{z^2}{2}} dz.$$

Відомо, що при довільному a

$$\int_{-\infty - ia}^{\infty - ia} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}.$$

Отже характеристична функція величини ξ дорівнює

$$f\left(t\right)=e^{ita-\frac{t^{2}\sigma^{2}}{2}}.$$

Обчислимо моменти випадкової величини ξ, яка підпорядкована нормальному закону, причому обмежимося тут випадком нормально розподіленої випадкової величини для якої $a=0$, $\sigma=1$.

Користуючись формулою Лейбніца для похідних добутку, маємо

$$f^{(n)}\left(t\right)=\left(e^{-\frac{t^{2}}{2}}\right)^{(n)}=\left(-te^{-\frac{t^{2}}{2}}\right)^{(n-1)}=-t\left(e^{-\frac{t^{2}}{2}}\right)^{(n-1)}-$$

$$-\left(n-1\right)\left(e^{-\frac{t^{2}}{2}}\right)^{(n-2)}=-tf^{(n-1)}\left(t\right)-\left(n-1\right)f^{(n-2)}\left(t\right).$$

Поклавши $t=0$, отримаємо

$$f^{(n)}\left(0\right)=-\left(n-1\right)f^{(n-2)}\left(0\right).$$

Отже,

$$\mathsf{E}\xi^{n}=\frac{1}{i^{n}}f^{(n)}\left(0\right)=-\frac{n-1}{i^{n}}f^{(n-2)}\left(0\right)=$$

$$=\frac{n-1}{i^{n-2}}f^{(n-2)}\left(0\right)=\left(n-1\right)\mathsf{E}\xi^{n-2}.$$

Беручи до уваги те, що $\mathsf{E}\xi^{0}=1$, маємо:

$$\mathsf{E}\xi^{2}=1,\quad \mathsf{E}\xi^{4}=1\cdot 3,\quad \mathsf{E}\xi^{6}=1\cdot 3\cdot 5,$$

і взагалі для парних n

$$\mathsf{E}\xi^{n}=1\cdot 3\cdot 5\cdots\left(n-1\right).$$

Так само, маючи на увазі те, що $\mathsf{E}\xi=0$, дістаємо для n непарних $\mathsf{E}\xi^{n}=0$.

Приклад 6.2. Знайдемо характеристичну функцію випадкової величини, розподіленої за законом Пуассона. Величина ξ набуває лише невід'ємні цілі значення, причому

$$P\left\{\xi=k\right\}=\frac{\lambda^{k}}{k!}e^{-\lambda}\quad k=0,1,2,\ldots,$$

де $\lambda>0$ – стала.

Характеристична функція випадкової величини ξ дорівнює

$$f(t) = Me^{it\xi} = \sum_{k=0}^{\infty} e^{itk} P\{\xi = k\} = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} =$$
$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{it}\right)^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda\left(e^{it}-1\right)}.$$

Згідно із (6.6) знаходимо, що

$$\mathsf{E}\xi = \frac{1}{i}\psi'(0) = \lambda, \quad D\xi = -\psi''(0) = \lambda.$$

Ці рівності були вже знайдені раніше безпосередньо.

Приклад 6.3. Випадкова величина ξ рівномірно розподілена в інтервалі $(-a, a)$. Характеристична функція величини ξ дорівнює

$$f(t) = \int_{-a}^{a} e^{itx} \frac{dx}{2a} = \frac{1}{2a} \frac{e^{ita} - e^{-ita}}{it} = \frac{\sin at}{at}.$$

Приклад 6.4. Знайти характеристичну функцію величини μ, що дорівнює кількості появ події A в n незалежних випробуваннях, у кожному з яких імовірність появи події A дорівнює p. Величина μ може бути подана як сума

$$\mu = \mu_1 + \mu_2 + \ldots \mu_n$$

n незалежних випадкових величин, кожна з яких набуває лише 2 значення: 0 та 1 – відповідно з імовірностями $q = 1 - p$ та p. Величина μ_k набуває значення 1, коли подія A відбувається при k-му випробуванні, і значення 0, коли подія A при k-му випробуванні не відбувається. Характеристична функція величини μ_k дорівнює

$$f_k(t) = Me^{it\mu k} = e^{it\cdot 0} \cdot q + e^{it\cdot 1} \cdot p = q + pe^{it}.$$

Згідно з теоремою 6.7 характеристична функція величини μ дорівнює

$$f(t) = \prod_{k=1}^{n} f_k(t) = \left(q + pe^{it}\right)^n.$$

Знайдемо ще характеристичну функцію величини

$$\eta = \frac{\mu - np}{\sqrt{npq}}.$$

За теоремою 6.2 вона дорівнює

$$f_\eta(t) = e^{-it\sqrt{\frac{np}{q}}} f\left(\frac{t}{\sqrt{npq}}\right) = e^{-it\sqrt{\frac{np}{q}}} \left(q + pe^{i\frac{t}{\sqrt{npq}}}\right)^n =$$

$$= \left(qe^{-it\sqrt{\frac{p}{nq}}} + pe^{it\sqrt{\frac{q}{np}}}\right)^n.$$

Приклад 6.5. Нехай випадкова величина ξ має гама розподіл з параметрами (α, β), тобто $\xi \simeq \Gamma(\alpha, \beta)$. Знайдемо характеристичну функцію випадкової величини ξ. Випадкова невід'ємна величина ξ має гама розподіл з параметрами (α, β), якщо її щільність розподілу має вигляд

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0. \end{cases}$$

Характеристична функція величини ξ дорівнюює

$$f(t) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} e^{itx} dx = \frac{\beta^\alpha}{(\beta - it)^\alpha}.$$

Знаючи характеристичну функцію випадкової величини $\xi \simeq \Gamma(\alpha, \beta)$ можемо записати характеристичну функцію випадкової величини розподіленої за показниковим розподілом $\xi \simeq Exp(\lambda)$. Маємо

$$f(t) = \frac{\lambda}{(\lambda - it)}.$$

Можемо записати характеристичну функцію випадкової величини розподіленої за χ_n^2 розподілом. Маємо

$$f(t) = \frac{1}{(1 - 2it)^{\frac{n}{2}}}.$$

Можемо також записати характеристичну функцію випадкової величини розподіленої за розподілом Ерланга. Маємо

$$f(t) = \frac{\lambda^n}{(\lambda - it)^n}.$$

Приклад 6.6. Характеристичні функції $f(t)$ задовольняють рівність

$$f(-t) = \overline{f(t)}.$$

Справді

$$f(-t) = \int\limits_{-\infty}^{\infty} e^{-itx} dF(x) = \overline{\int\limits_{-\infty}^{\infty} e^{itx} dF(x)} = \overline{f(t)}.$$

Приклад 6.7. Знайти характеристичну функцію випадкової величини ξ, розподіленої за законом Коші:

$$p(x) = \frac{1}{\pi} \frac{a}{a^2 + x^2}.$$

Маємо:

$$f(t) = \frac{a}{\pi} \int\limits_{-\infty}^{\infty} \frac{e^{itx}}{a^2 + x^2} dx = \frac{a}{\pi} \int\limits_{-\infty}^{\infty} \frac{\cos tx}{1 + x^2} dx.$$

Функція $f(t)$ парна. Знайдемо $f(t)$ при $t > 0$. Похідна

$$f'(t) = \frac{a}{\pi} \int\limits_{-\infty}^{\infty} \frac{-x \sin tx}{1 + x^2} dx.$$

Скористаємося відодим інтегралом (Діріхле)

$$\pi = \int\limits_{-\infty}^{\infty} \frac{\sin tx}{x} dx, \quad t > 0.$$

Запишемо його у вигляді

$$a = \frac{a}{\pi} \int\limits_{-\infty}^{\infty} \frac{\sin tx}{x} dx.$$

Тоді

$$f'(t) + a = \frac{a}{\pi} \int\limits_{-\infty}^{\infty} \frac{a^2}{a^2 + x^2} \frac{\sin tx}{x} dx.$$

Дифференціювання по t дасть нам рівняння для $f(t)$

$$f''(t) = a f(t).$$

Звідки знаходимо характеристичну функцію

$$f(t) = e^{-a|t|}.$$

Аналогічно отримаємо, що характеристична функція випадкової величини ξ, розподіленої за законом Коші

$$p\,(x) = \frac{a}{\pi}\,\frac{1}{a^2 + (x-\theta)^2}$$

дорівнює

$$f(t) = e^{it\theta - a|t|}.$$

Зауважимо, що для розподілу, який ми розглядаємо, не існує скінченних математичного сподівання і дисперсії. Це можна побачити і безпосередньо: характеристична функція недиференційовна при $t = 0$.

6.2 Формула обернення та теорема єдиності

За функцією розподілу випадкової величини ξ можна знайти її характеристичну функцію. Для нас важливо, що має місце також обернене твердження: за характеристичною функцією функція розподілу визначається однозначно.

Теорема 6.8. *Нехай $f(t)$ і $F(x)$ – характеристична функція і функція розподілу випадкової величини ξ. Якщо x_1 і x_2 – точки неперервності функції $F(x)$, то*

$$F\,(x_2) - F\,(x_1) = \lim_{c\to+\infty} \frac{1}{2\pi} \int\limits_{-c}^{c} \frac{e^{-itx_1} - e^{-itx_2}}{it} f\,(t)\,dt. \qquad (6.7)$$

Доведення. З означення характеристичної функції випливає рівність

$$I_c = \frac{1}{2\pi} \int\limits_{-c}^{c} \frac{e^{-itx_1} - e^{-itx_2}}{it} f\,(t)\,dt =$$

$$= \frac{1}{2\pi} \int\limits_{-c}^{c} \int\limits_{-\infty}^{\infty} \frac{1}{it} \left[e^{it(z-x_1)} - e^{it(z-x_2)}\right] dF\,(z)\,dt.$$

В останньому інтегралі можна змінити порядок інтегрування, оскільки відносно z інтеграл абсолютно збігається, а відносно t границі інтегрування скінченні. Таким чином,

$$I_c = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\left[\int\limits_{-c}^{c}\frac{e^{it(z-x_1)} - e^{it(z-x_2)}}{it}dt\right]dF(z) =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\left[\int\limits_{0}^{c}\frac{1}{it}\left\{e^{it(z-x_1)} - e^{-it(z-x_1)} - e^{it(z-x_2)} + e^{-it(z-x_2)}\right\}dt\right]dF(z) =$$

$$= \frac{1}{\pi}\int\limits_{-\infty}^{\infty}\int\limits_{0}^{c}\left[\frac{\sin t(z-x_1)}{t} - \frac{\sin t(z-x_2)}{t}\right]dtdF(z).$$

Пригадаємо із курсу математичного аналізу, що при $c \to \infty$

$$\frac{1}{\pi}\int\limits_{0}^{c}\frac{\sin at}{t}dt \to \begin{cases} \frac{1}{2}, & a > 0, \\ -\frac{1}{2}, & a < 0, \end{cases} \tag{6.8}$$

і ця збіжність рівномірна відносно a в кожному інтервалі $a > \delta > 0$ (відповідно $a < -\delta$) і що при $|a| \le \delta$ за всіх досить великих c

$$\left|\frac{1}{\pi}\int\limits_{0}^{c}\frac{\sin at}{t}dt\right| < 1. \tag{6.9}$$

Вважатимемо, що $x_2 > x_1$ і зобразимо інтеграл I_c у вигляді такої суми:

$$I_c = \int\limits_{-\infty}^{x_1-\delta} + \int\limits_{x_1-\delta}^{x_1+\delta} + \int\limits_{x_1+\delta}^{x_2-\delta} + \int\limits_{x_2-\delta}^{x_2+\delta} + \int\limits_{x_2+\delta}^{\infty}\psi(c,z;x_1,x_2)\,dF(z),$$

де для скорочення позначимо

$$\psi(c,z;x_1,x_2) = \frac{1}{\pi}\int\limits_{0}^{c}\left\{\frac{\sin t(z-x_1)}{t} - \frac{\sin t(z-x_2)}{t}\right\}dt,$$

і $\delta > 0$ підібрано так, що $x_1 + \delta < x_2 - \delta$. Із того, що в інтервалі $-\infty < z < x_1 - \delta$ мають місце нерівності $z - x_1 < -\delta$ та $z - x_2 < -\delta$, на підставі (6.8) виводимо, що при $c \to \infty$

$$\int\limits_{-\infty}^{x_1-\delta} \psi\left(c, z; x_1, x_2\right) dF\left(z\right) \to 0.$$

Із того, що $z - x_1 > \delta$ і $z - x_2 > \delta$ при $x_2 + \delta < z < +\infty$ із (6.8) випливає, що при $c \to \infty$

$$\int\limits_{x_2+\delta}^{\infty} \psi\left(c, z; x_1, x_2\right) dF\left(z\right) \to 0.$$

Через те, що в інтервалі $x_1 + \delta < z < x_2 - \delta$ мають місце нерівності $z - x_1 > \delta$ та $z - x_2 < -\delta$, то згідно з (6.8) при $c \to \infty$

$$\int\limits_{x_1+\delta}^{x_2-\delta} \psi\left(c, z; x_1, x_2\right) dF\left(z\right) \to \int\limits_{x_1+\delta}^{x_2-\delta} dF\left(z\right) = F\left(x_2 - \delta\right) - F\left(x_1 + \delta\right).$$

Нарешті, на підставі (6.9) можна скористатися такими оцінками: за всіх досить великих с

$$\left|\int\limits_{x_1-\delta}^{x_1+\delta} \psi\left(c, z; x_1, x_2\right) dF\left(z\right)\right| < 2\int\limits_{x_1-\delta}^{x_1+\delta} dF\left(z\right) = 2\left[F\left(x_1 + \delta\right) - F\left(x_1 - \delta\right)\right],$$

$$\left|\int\limits_{x_2-\delta}^{x_2+\delta} \psi\left(c, z; x_1, x_2\right) dF\left(z\right)\right| < 2\int\limits_{x_2-\delta}^{x_2+\delta} dF\left(z\right) = 2\left[F\left(x_2 + \delta\right) - F\left(x_2 - \delta\right)\right].$$

Таким чином, знаходимо, що при довільному $\delta > 0$

$$\lim_{c\to\infty} I_c = F\left(x_2 - \delta\right) - F\left(x_1 + \delta\right) + R\left(\delta, x_1, x_2\right),$$

де

$$\left|R\left(\delta, x_1, x_2\right)\right| < 2\left\{F\left(x_1 + \delta\right) - F\left(x_1 - \delta\right) + F\left(x_2 + \delta\right) - F\left(x_2 - \delta\right)\right\}.$$

Припустимо тепер, що $\delta \to 0$, тоді із того, що $_1$ та $_2$ є точками неперервності функції $F(x)$, випливають рівності

$$\lim_{\delta\to 0} F\left(x_1 + \delta\right) = \lim_{\delta\to 0} F\left(x_1 - \delta\right) = F\left(x_1\right),$$

$$\lim_{\delta\to 0} F\left(x_2+\delta\right) = \lim_{\delta\to 0} F\left(x_2-\delta\right) = F\left(x_2\right),$$

а оскільки I_c не залежить від δ, то має місце рівність

$$\lim_{c\to\infty} I_c = F\left(x_2\right) - F\left(x_1\right).$$

Очевидно, що це співвідношення доводить нашу теорему. $\square$

Рівність (6.7) має назву формули обернення. Використаємо цю формулу для доведення такого важливого твердження.

Теорема 6.9. *Функція розподілу однозначно визначаеться своєю характеристичною функцією.*

Доведення. Припустимо протилежне, що дві різні функції розподілу $F_1(x)$ та $F_2(x)$ мають одну й ту ж характеристичну функцію $f(t)$. Тоді згідно з формулою обернення, якщо x_1 та x_2 є точками неперервності як функції $F_1(x)$ так і $F_2(x)$, має місце рівність

$$F_1\left(x_2\right) - F_1\left(x_1\right) = F_2\left(x_2\right) - F_2\left(x_1\right) =$$

$$= \frac{1}{2\pi}\lim_{c\to\infty}\int\limits_{-c}^{c}\frac{e^{-itx_1}-e^{-itx_2}}{it}f\left(t\right)dt. \tag{6.10}$$

Оскільки точок розриву функцій $F_1(x)$ та $F_2(x)$ може бути щонайбільше зліченна множина, то існують як завгодно великі за абсолютною величиною точки x_1 та x_2, які є одночасно точками неперервності функцій $F_1(x)$ та $F_2(x)$. Припустимо тепер, що $x_1 \to -\infty$, при цьому $F_1\left(x_1\right) \to 0$ і $F_2\left(x_1\right) \to 0$, із рівності (6.10) дістанемо, що у всіх точках x_2 неперервності функцій $F_1(x)$ та $F_2(x)$

$$F_1\left(x_2\right) = F_2\left(x_2\right).$$

Ця рівність доводить нашу теорему. $\square$

Застосуємо останню теорему до доведення таких тверджень.

Приклад 6.8. Якщо незалежні випадкові величини ξ_1 та ξ_2 розподілені за нормальним законом, то їхня сума $\xi = \xi_1 + \xi_2$ також розподілена за нормальним законом. Справді, якщо

$$\mathsf{E}\xi_1 = a_1,\ \mathsf{E}\xi_2 = a_2,\ D\xi_1 = \sigma_1^2,\ D\xi_2 = \sigma_2^2,$$

то характеристичні функції величин ξ_1 та ξ_2 дорівнюють відповідно:

$$f_1(t) = e^{ia_1 t - \frac{1}{2}\sigma_1^2 t^2}, \quad f_2(t) = e^{ia_2 t - \frac{1}{2}\sigma_2^2 t^2}.$$

Характеристична функція $f(t)$ суми $\xi = \xi_1 + \xi_2$ дорівнює

$$f(t) = f_1(t) f_2(t) = e^{it(a_1+a_2) - \frac{1}{2}\left(\sigma_1^2 + \sigma_2^2\right)t^2}.$$

Це характеристична функція нормального закону з математичним сподіванням $a = a_1 + a_2$ і дисперсією $\sigma^2 = \sigma_1^2 + \sigma_2^2$. На підставі теореми єдиності з отриманої рівності виводимо, що функція розподілу випадкової величини ξ є нормальною функцією розподілу.

Приклад 6.9. Незалежні випадкові величини ξ_1 та ξ_2 набувають цілі невід'ємні значення, причому

$$P\{\xi_1 = k\} = \frac{\lambda_1^k e^{-\lambda_1}}{k!}$$

та

$$P\{\xi_2 = k\} = \frac{\lambda_2^k e^{-\lambda_2}}{k!}.$$

Доведемо, що випадкова величина $\zeta = \xi_1 + \xi_2$ розподілена за законом Пуассона з параметром $\lambda = \lambda_1 + \lambda_2$. Згідно з прикладом попереднього параграфа характеристичні функції величин ξ_1 та ξ_2 дорівнюють відповідно

$$f_1(t) = e^{\lambda_1\left(e^{it} - 1\right)}, \quad f_2(t) = e^{\lambda_2\left(e^{it} - 1\right)}.$$

Характеристична функція суми $\zeta = \xi_1 + \xi_2$ дорівнює

$$f(t) = f_1(t) f_2(t) = e^{(\lambda_1 + \lambda_2)\left(e^{it} - 1\right)},$$

тобто є характеристичною функцією певного закону Пуассона. Згідно з теоремою єдиності функція розподілу величини ζ є законом Пуассона з параметром $\lambda = \lambda_1 + \lambda_2$. Для нього

$$P\{\xi = k\} = \frac{(\lambda_1 + \lambda_2)^k e^{-(\lambda_1 + \lambda_2)}}{k!}, \quad k = 0, 1, 2, \ldots$$

Вірне і зворотнє твердження. Якщо сума двох незалежних випадкових величин розподілена за законои Пуассона, то кожний доданок теж розподілений за законои Пуассона.

Приклад 6.10. Нехай незалежні випадкові величини ξ_1 та ξ_2 мають гама розподіли з параметрами (α_1, β) та (α_2, β), тобто $\xi_1 \simeq \Gamma(\alpha_1, \beta)$ та $\xi_2 \simeq \Gamma(\alpha_2, \beta)$.

Випадкова невід'ємна величина ξ має гама розподіл з параметрами (α, β) (позначення $\xi \simeq \Gamma(\alpha, \beta)$), якщо її щільність розподілу має вигляд

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0. \end{cases}$$

Покажемо, що випадкова величина $\eta = \xi_1 + \xi_2$ має гама розподіл з параметрами $(\alpha_1 + \alpha_2, \beta)$, тобто $\xi_1 + \xi_2 \simeq \Gamma(\alpha_1 + \alpha_2, \beta)$.

Дійсно, характеристичні функції величин ξ_1 та ξ_2 дорівнюють відповідно

$$f_1(t) = \frac{\beta^{\alpha_1}}{(\beta - it)^{\alpha_1}}, \quad f_2(t) = \frac{\beta^{\alpha_2}}{(\beta - it)^{\alpha_2}}.$$

Характеристична функція суми $\zeta = \xi_1 + \xi_2$ дорівнює добутку характеристичних функцій

$$f(t) = f_1(t) \cdot f_2(t) = \frac{\beta^{\alpha_1}}{(\beta - it)^{\alpha_1}} \cdot \frac{\beta^{\alpha_2}}{(\beta - it)^{\alpha_2}} = \frac{\beta^{(\alpha_1 + \alpha_2)}}{(\beta - it)^{(\alpha_1 + \alpha_2)}},$$

тобто є характеристичною функцією певного гама розподілу. Згідно з теоремою єдиності випадкова величина $\zeta = \xi_1 + \xi_2$ має гама розподіл з параметрами $(\alpha_1 + \alpha_2, \beta)$, тобто $\xi_1 + \xi_2 \simeq \Gamma(\alpha_1 + \alpha_2, \beta)$.

Приклад 6.11. Характеристична функція випадкової величини ξ дійсна тоді й тільки тоді, коли функція розподілу величини ξ симетрична, тобто коли за довільних x для неї має місце рівність $F(x) = 1 - F(-x + 0)$. Якщо функція розподілу симетрична, то легко бачити, що її характеристична функція дійсна. Справді, у цьому випадку

$$\begin{aligned} f(t) &= \int_{-\infty}^{\infty} e^{itx} dF(x) = \int_{-\infty}^{0} e^{itx} dF(x) + \int_{0}^{\infty} e^{itx} dF(x) = \\ &= \int_{0}^{\infty} e^{itx} d[1 - F(-x)] + \int_{0}^{\infty} e^{itx} dF(x) = \\ &= \int_{0}^{\infty} \left(e^{itx} + e^{-itx}\right) dF(x) = 2 \int_{0}^{\infty} \cos tx dF(x). \end{aligned}$$

Для доведення оберненого твердження розглянемо випадкову величину $\eta = -\xi$. Функція розподілу величини η дорівнює

$$\Phi(x) = P\{\eta < x\} = P\{\xi > -x\} = 1 - F(-x + 0).$$

Характеристичні функції величин ξ та η зв'язані співвідношенням:

$$\varphi(t) = Me^{it\eta} = Me^{-it\xi} = \overline{Me^{-it\xi}} = \overline{f(t)}.$$

Через те, що $f(t)$ – дійсна, $\overline{f(t)} = f(t)$ і, отже, $\varphi(t) = f(t)$. Із теореми єдиності робимо висновок, що функції розподілу величин ξ та η збігаються, тобто

$$F(x) = 1 - F(-x+0),$$

що й треба було довести.

6.3 Теореми Хеллі

Розглянемо дві теореми чисто аналітичного характеру – перша і друга теореми Хеллі (Helly).

Означення 6.3. Кажуть, що послідовність неспадних функцій

$$F_1(x), F_2(x), \ldots F_n(x), \ldots \tag{6.11}$$

збігається до неспадної функції $F(x)$ в основному, якщо при $n \to \infty$ вона збігається до неї в кожній точці неперервності $F(x)$.

Далі завжди будемо вважати, що функції $F_n(x)$ задовольняють додаткову умову $F_n(-\infty) = 0$.

Відмітимо відразу ж, що для збіжності в основному досить, щоб послідовність (6.11) збігалася до функції $F(x)$ на якійсь всюди щільній множині D. Справді, нехай x – довільна точка і нехай x' та x'' – якісь дві точки множини D, такі, що $x' \le x \le x''$. Тоді й

$$F_n(x') \le F_n(x) \le F_n(x'').$$

Отже,

$$\lim_{n\to\infty} F_n(x') = \varliminf_{n\to\infty} F_n(x') \le \varliminf_{n\to\infty} F_n(x) \le$$
$$\le \varlimsup_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x'') = \lim_{n\to\infty} F_n(x'').$$

Оскільки за припущенням

$$\lim_{n\to\infty} F_n(x') = F(x'), \quad \lim_{n\to\infty} F_n(x'') = F(x''),$$

то й

$$F(x') \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x'').$$

Але середні члени в цих нерівностях не залежать від x' та x'', і тому також

$$F\left(x-0\right) \leq \varliminf_{n\to\infty} F_n\left(x\right) \leq \varlimsup_{n\to\infty} F_n\left(x\right) \leq F\left(x+0\right).$$

Якщо тепер функція $F\left(x\right)$ в точці х неперервна, то

$$F\left(x-0\right) = F\left(x\right) = F\left(x+0\right).$$

Тому в точках неперервності останні нерівності дають

$$\varliminf_{n\to\infty} F_n\left(x\right) = \varlimsup_{n\to\infty} F_n\left(x\right) = \lim_{n\to\infty} F_n\left(x\right) = F\left(x\right),$$

що й треба було встановити.

Теорема 6.10. (Перша теорема Хеллі.) *Перша теорема Хеллі. Кожна послідовність обмежених у сукупності неспадних функцій*

$$F_1\left(x\right), F_2\left(x\right), \ldots F_n\left(x\right), \ldots \tag{6.12}$$

містить щонайменше одну підпослідовність

$$F_{n_1}\left(x\right), F_{n_2}\left(x\right), \ldots F_{n_k}\left(x\right), \ldots$$

яка збігається в основному до певної неспадної функції $F\left(x\right)$.

Доведення. Нехай D – зліченна всюди щільна множина точок $x_1', x_2', \ldots, x_n', \ldots$ Візьмемо значення функції послідовності (6.12) в точці x_1':

$$F_1\left({x'}_1\right), F_2\left({x'}_1\right), \ldots F_n\left({x'}_1\right), \ldots$$

Оскільки множина цих значень обмежена, то вона містить щонайменше одну послідовність, яка збігається до певної границі $F\left({x'}_1\right)$. Інакше кажучи, послідовність (6.12) містить щонайменше одну підпослідовність

$$F_{11}\left(x\right), F_{12}\left(x\right), \ldots F_{1n}\left(x\right), \ldots \tag{6.13}$$

для якої

$$\lim_{n\to\infty} F_{1n}\left(x_1'\right) = F\left(x_1'\right).$$

Далі візьмемо значення всіх функцій послідовності (6.13) в точці ${x'}_2$:

$$F_{11}\left(x_2'\right), F_{12}\left(x_2'\right), \ldots F_{1n}\left(x_2'\right), \ldots$$

Оскільки множина цих значень також обмежена, то й вона містить щонайменше одну підпослідовність, яка збігається до певного граничного значення $F\left(x'_2\right)$, тобто послідовність (6.13) містить щонайменше одну підпослідовність

$$F_{21}\left(x\right), F_{22}\left(x\right), \ldots F_{2n}\left(x\right), \ldots \tag{6.14}$$

для якої

$$\lim_{n \to \infty} F_{2n}\left(x'_2\right) = F\left(x'_2\right).$$

Продовжуючи таке виділення підпослідовностей з одержаних раніше послідовностей, для кожної точки x'_k множини D дістанемо послідовність

$$F_{k1}\left(x\right), F_{k2}\left(x\right), \ldots F_{kn}\left(x\right), \ldots \tag{6.15}$$

для якої

$$\lim_{n \to \infty} F_{kn}\left(x\right) = F\left(x'_k\right).$$

Утворимо тепер діагональну послідовність

$$F_{11}\left(x\right), F_{22}\left(x\right), \ldots F_{nn}\left(x\right), \ldots \tag{6.16}$$

Вона виділена із послідовності (6.13); тому

$$\lim_{n \to \infty} F_{nn}\left(x'_1\right) = F\left(x'_1\right).$$

Уся діагональна послідовність, за винятком лише першого члена $F_{11}\left(x\right)$, виділена із послідовності (6.14); тому

$$\lim_{n \to \infty} F_{nn}\left(x'_2\right) = F\left(x'_2\right).$$

Взагалі вся діагональна послідовність, за винятком перших $(k-1)$ членів

$$F_{11}\left(x\right), F_{22}\left(x\right), \ldots F_{k-1,k-1}\left(x\right),$$

виділена із послідовності (6.15); тому

$$\lim_{n \to \infty} F_{nn}\left(x'_k\right) = F\left(x'_k\right)$$

при кожному k.

Цей результат можна сформулювати так: послідовність (6.12) містить щонайменше одну підпослідовність

$$F_{n_1}\left(x\right), F_{n_2}\left(x\right), \ldots F_{n_k}\left(x\right), \ldots$$

яка для всіх точок x'_n множини D збігається до певної функції $F(x)$, визначеної на цій множині D. При цьому, через те що $F_{n_k}(x)$ – неспадні й рівномірно обмежені, то, очевидно, і функція $F(x)$ буде неспадною й обмеженою. Тепер ясно, що функцію $F(x)$, визначену на множині D, можна продовжити так, що вона буде задана на всій прямій, $-\infty < x < \infty$, лишаючись неспадною й обмеженою: для цього досить покласти в довільній точці x, яка не належить до множини D,

$$F(x) = \lim_{x' \to x} F(x'),$$

де x' належить до множини і лежить ліворуч від точки x. Послідовність (6.16) збігається до цієї функції на скрізь щільній множині D. Тому вона збігається до неї в основному, що й доводить теорему. $\square$

Теорема 6.11. (Друга теорема Хеллі.) *Нехай $f(x)$ – неперервна функція і послідовність*

$$F_1(x), F_2(x), \ldots F_n(x), \ldots$$

неспадних, обмежених у сукупності функцій збігається в основному до функції $F(x)$ на певному скінченному інтервалі $a \le x \le b$, де a та b – точки неперервності функції $F(x)$, тоді

$$\lim_{n \to \infty} \int_a^b f(x)\, dF_n(x) = \int_a^b f(x)\, dF(x).$$

Доведення. Із неперервності функції $f(x)$ випливає, що яке б не було додатне стале ε, знайдеться розділ інтервалу $a \le x \le b$ точками поділу $a = x_0, x_1, \ldots, x_N = b$ на інтервали довжини $\Delta x_0, \Delta x_1, \ldots, \Delta x_{N-1}$ такий, що в кожному інтервалі $x_k \le x \le x_{k+1}$ буде виконуватися нерівність

$$|f(x) - f(x_k)| < \varepsilon.$$

Користуючись цією обставиною, можна ввести допоміжну функцію $f_\varepsilon(x)$, яка набуває лише скінченну кількість значень, а саме – функцію, визначену рівностями

$$f_\varepsilon(x) = f(x_k)$$

при

$$x_k \le x < x_{k+1},$$

для якої в усьому інтервалі

$$a \le x \le b$$

виконуватиметься нерівність

$$|f(x) - f(x_k)| < \varepsilon. \tag{6.17}$$

При цьому можна, очевидно, наперед вибрати розбиття $x_0, x_1, \ldots, x_N$ так, щоб усі вони були точками неперервності функції $F(x)$. Тоді в силу збіжності в основному до функції $F(x)$ функцій $F_1(x), F_2(x), \ldots$ за досить великих n в усіх цих точках поділу будуть виконуватися нерівності

$$|F(x_k) - F_n(x_k)| < \frac{\varepsilon}{MN}, \tag{6.18}$$

де M – максимум модуля $|f(x)|$ в інтервалі $a \le x \le b$. Очевидно, що

$$\begin{gathered}
\left| \int_a^b f(x)\, dF(x) - \int_a^b f(x)\, dF_n(x) \right| \le \\
\le \left| \int_a^b f(x)\, dF(x) - \int_a^b f_\varepsilon(x)\, dF(x) \right| + \\
+ \left| \int_a^b f_\varepsilon(x)\, dF(x) - \int_a^b f_\varepsilon(x)\, dF_n(x) \right| + \\
+ \left| \int_a^b f_\varepsilon(x)\, dF_n(x) - \int_a^b f(x)\, dF_n(x) \right|. \qquad (6.19)
\end{gathered}$$

Неважко бачити, що в цій нерівності перший доданок правої частини не перевищує

$$\varepsilon\left[F(b) - F(a)\right],$$

а третій не перевищує

$$\varepsilon\left[F_n(b) - F_n(a)\right];$$

це випливає з нерівності (6.17). Що ж до другого доданку, він, як легко перевірити, дорівнює

$$\left|\sum_{k=0}^{N-1} f(x_k)\left[F(x_{k+1})-F(x_k)\right]-\sum_{k=0}^{N-1} f(x_k)\left[F_n(x_{k+1})-F_n(x_k)\right]\right|=$$

$$=\left|\sum_{k=0}^{N-1} f(x_k)\left[F(x_{k+1})-F_n(x_{k+1})\right]-\sum_{k=0}^{N-1} f(x_k)\left[F(x_k)-F_n(x_k)\right]\right|$$

і, отже, за досить великих n він не перевищує 2ε; це випливає з нерівності (6.18). У силу рівномірної обмеженості функції $F_n(x)$ сума

$$\varepsilon\left[F(b)-F(a)\right]+\varepsilon\left[F_n(b)-F_n(a)\right]+2\varepsilon$$

може бути зроблена як завгодно малою разом з ε. Тому за досить великих n буде як завгодно малою і ліва частина нерівності (6.19), що й доводить нашу теорему. □

Теорема 6.12. (Узагальнення другої теореми Хеллі.) *Нехай функція $f(x)$ неперервна і обмежена на всій прямій $-\infty < x < +\infty$ і послідовність*

$$F_1(x), F_2(x), \ldots F_n(x), \ldots$$

функції розподілу випадкових величин збігається в основному до функції розподілу $F(x)$, тоді

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f(x)\,dF_n(x)=\int_{-\infty}^{\infty} f(x)\,dF(x).$$

Доведення. Візьмемо довільні число $<0, B>0$ і покладемо

$$I_1=\left|\int_{-\infty}^{A} f(x)\,dF(x)-\int_{-\infty}^{A} f(x)\,dF_n(x)\right|,$$

$$I_2=\left|\int_{A}^{B} f(x)\,dF(x)-\int_{A}^{B} f(x)\,dF_n(x)\right|,$$

$$I_3=\left|\int_{B}^{\infty} f(x)\,dF(x)-\int_{B}^{\infty} f(x)\,dF_n(x)\right|.$$

Тоді, очевидно, що

$$\left| \int\limits_{-\infty}^{\infty} f(x)\, dF(x) - \int\limits_{-\infty}^{\infty} f(x)\, dF_n(x) \right| \leq I_1 + I_2 + I_3.$$

Величини I_1 та I_3 можуть бути зроблені як завгодно малими, якщо підібрати A і B досить великими за абсолютною величиною і до того ж такими, щоб точки A і B були точками неперервності функції $F(x)$ і взяти потім досить велике значення n. Справді, нехай M – верхня границя модуля $|f(x)|$ на всій прямій $-\infty \leq x \leq \infty$. Тоді

$$I_1 \leq M \left[F(A) + F_n(A) \right],$$

$$I_3 \leq M \left[(1 - F(B)) + (1 - F_n(B)) \right].$$

Але

$$\lim_{A \to -\infty} F(A) = 0, \quad \lim_{B \to +\infty} F(B) = 1,$$

і на підставі збіжності в основному функції $F_n(x)$ до функції $F(x)$

$$\lim_{n \to \infty} F(A) = F(A), \quad \lim_{n \to \infty} F(B) = F(B).$$

Це й доводить твердження про I_1 та I_3. Величину ж I_2 можна зробити як завгодно малою за будь-якого вибору A та B, коли точки A і B є точками неперервності функції $F(x)$, якщо взяти досить велике n. Справді, із неперервності функції $f(x)$ на всій прямій $-\infty \leq x \leq \infty$ випливає, що вона неперервна на інтервалі $A \leq x \leq B$; тому тут може бути застосована друга теорема Хеллі для обмеженого інтервалу. Ясно, що всі вимоги, які ми висуваємо до підбору A і B, а потім n, сумісні, і тому сума $I_1 + I_2 + I_3$ може бути зроблена як завгодно малою. Тим самим узагальнення другої теореми Хеллі повністю доведено. □

6.4 Граничні теореми для характеристичних функцій

Найважливішими з погляду застосувань характеристичних функцій до виведення асимптотичних формул теорії ймовірностей є дві граничні теореми – пряма та обернена. На підставі цих теорем можемо стверджувати, що відповідність, яка встановлена між функціями розподілу та характеристичними функціями, не тільки взаємно однозначна, а й неперервна.

Теорема 6.13. (Пряма гранична теорема.) *Якщо послідовність*

$$F_1(x), F_2(x), \dots F_n(x), \dots$$

функції розподілу випадкових величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ збігається в основному до функції розподілу $F(x)$ випадкової величини ξ, то і послідовність характеристичних функцій

$$f_1(t), f_2(t), \dots, f_n(t), \dots$$

випадкових величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ збігається до характеристичної функції $f(t)$ випадкової величини ξ, причому ця збіжність рівномірна на кожному скінченному інтервалі зміни t.

Доведення. Доведення. Оскільки

$$f_n(t) = \int\limits_{-\infty}^{\infty} e^{itx} dF_n(x), \quad f(t) = \int\limits_{-\infty}^{\infty} e^{itx} dF(x),$$

$$e^{itx} = \cos tx + i \sin tx,$$

а функції $\cos tx$ та $\sin tx$ на кожному t неперервні й обмежені на всій прямій $-\infty \leq t \leq \infty$, то з другої теореми Хеллі випливає, що при $n \to \infty$

$$f_n(t) \to f(t).$$

Твердження, що ця збіжність рівномірна в кожному скінченному інтервалі зміни t, перевіряється за допомогою таких самих міркувань, як при доведенні другої теореми Хеллі. □

Теорема 6.14. (Обернена гранична теорема.) *Якщо послідовність*

$$f_1(t), f_2(t), \dots f_n(t), \dots$$

характеристичних функцій випадкових величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ збігається до неперервної функції $f(t)$, то послідовність

$$F_1(x), F_2(x), \dots F_n(x), \dots \tag{6.20}$$

функції розподілу величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ збігається в основному до функції розподілу $F(x)$ певної випадкової величини ξ (у силу прямої граничної теореми $f(t)$ є характеристичною функцією цієї величини ξ).

Доведення. На підставі першої теореми Хеллі робимо висновок, що послідовність $\{F_n(x)\}$ обов'язково містить послідовність

$$F_{n_1}(x), F_{n_2}(x), \ldots F_{n_k}(x), \ldots$$

яка збігається в основному до певної неспадної функції $F(x)$. При цьому зрозуміло, що цю останню можна підібрати так, щоб у кожній точці x виконувалася рівність

$$\lim_{x' \to x-0} F(x') = F(x).$$

Ця функція $F(x)$ взагалі може не бути функцією розподілу деякої випадкової величини, оскільки для цього ще потрібно, щоб мали місце рівності $F(-\infty) = 0$ і $F(+\infty) = 1$. А якщо, наприклад,

$$F_n(x) = \begin{cases} 0, & x \le -n, \\ \frac{1}{2}, & -n < x \le n, \\ 1, & x > n, \end{cases}$$

то $F(x) = \frac{1}{2}$ тотожно, отже, $F(-\infty) = \frac{1}{2}$, а не $F(-\infty) = 0$. Проте в умовах нашої теореми буде $F(-\infty) = 0$, $F(+\infty) = 1$. Справді, нехай це не так, то беручи до уваги, що $F(-\infty) \ge 0$, $F(+\infty) \le 1$, матимемо

$$\delta = F(+\infty) - F(-\infty) < 1.$$

Візьмемо тоді яке-небудь додатне число ε, менше від $1-\delta$. За умовою нашої теореми послідовність характеристичних функцій $f_n(t)$ збігається до функції $f(t)$. Отже, $f(0) = 1$. А оскільки, крім того, функція $f(t)$ неперервна, можна вибрати досить мале додатне число τ так, що буде мати місце нерівність

$$\frac{1}{2\tau}\left|\int_{-\tau}^{\tau} f(t)\,dt\right| > 1 - \frac{\varepsilon}{2} > \delta + \frac{\varepsilon}{2} \tag{6.21}$$

Але в той самий час можна вибрати $X > \frac{4}{\tau\varepsilon}$ і таке велике k_0, щоб при було

$$\delta_k = F_{n_k}(X) - F_{n_k}(-X) < \delta + \frac{\varepsilon}{4}.$$

Через те, що $f_{n_k}(t)$ є характеристична функція ξ_{n_k}, маємо:

$$\int_{-\tau}^{\tau} f_{n_k}(t)\,dt = \int_{-\infty}^{\infty}\left(\int_{-\tau}^{\tau} e^{itx}dt\right) dF_{n_k}(x).$$

Інтеграл, записаний у правій частині цієї рівності, за змінною t можна оцінити так: з одного боку, оскільки $\left|e^{itx}\right| = 1$, то завжди

$$\left|\int\limits_{-\tau}^{\tau} e^{itx} dt\right| < 2\tau.$$

З іншого боку, оскільки

$$\int\limits_{-\tau}^{\tau} e^{itx} dt = \frac{e^{i\tau x} - e^{-i\tau x}}{ix} = 2\frac{\sin \tau x}{x},$$

а $|\sin \tau x| \leq 1$, то при $|x| > X$

$$\left|\int\limits_{-\tau}^{\tau} e^{itx} dt\right| < \frac{2}{X}$$

. Звідси, застосовуючи першу оцінку при $|x| \leq X$ і другу при $|x| > X$, маємо

$$\left|\int\limits_{-\tau}^{\tau} f_{n_k}(t)\, dt\right| \leq \left|\int\limits_{|x| \leq X} \int\limits_{-\tau}^{\tau} e^{itx} dt dF_{n_k}(x)\right| + \\ + \left|\int\limits_{|x| > X} \int\limits_{-\tau}^{\tau} e^{itx} dt dF_{n_k}(x)\right| < 2\tau\delta_k + \tfrac{2}{X}$$

і, отже,

$$\frac{1}{2\tau}\left|\int\limits_{-\tau}^{\tau} f_{n_k}(t)\, dt\right| < \delta + \frac{\varepsilon}{2}$$

. Нарешті, оскільки функції $f_{n_k}(t)$ обмежені в сукупності та збігаються до функції $f(t)$, можемо зробити граничний перехід у попередній нерівності й отримаємо

$$\frac{1}{2\tau}\left|\int\limits_{-\tau}^{\tau} f(t)\, dt\right| < \delta + \frac{\varepsilon}{2},$$

що суперечить нерівності (6.21). Отже, функція $F(x)$, до якої збігаються в основному функції $F_{n_k}(x)$, є функцією розподілу певної випадкової величини. У силу прямої граничної теореми це величина, характеристична функція якої є $f\left(t\right)$. Щоб закінчити доведення теореми, залишається показати, що і вся послідовність (6.20) збігається в основному до функції $F\left(x\right)$. Коли б це було не так, то знайшлася б точка с неперервності функції $F\left(x\right)$ і деяка підпослідовність послідовності $\{F_n\left(c\right)\}$, яка збігається в цій точці до числа, що не дорівнює $F\left(c\right)$. У цій же підпослідовності, у свою чергу, знайшлася б підпослідовність

$$F^*_{n_1}\left(x\right), F^*_{n_2}\left(x\right), \ldots F^*_{n_k}\left(x\right), \ldots$$

яка збігається в основному до певної функції $\{F^*\left(c\right)\}$; для цієї останньої, зокрема, мали б

$$F^*\left(c\right) \neq F\left(c\right).$$

Але, застосовуючи до послідовності $\left\{F^*_{n_k}\left(x\right)\right\}$ ті ж міркування, які було застосовано вище до послідовності $\{F_{n_k}\left(x\right)\}$, легко переконатися в тому, що її гранична функція $F^*\left(x\right)$ є функцією розподілу випадкової величини, яка має характеристичну функцію $f\left(t\right)$; отже, $F^*\left(x\right) = F\left(x\right)$ і, зокрема,

$$F^*\left(c\right) = F\left(c\right),$$

що суперечить тільки що записаній нерівності. □

Зауваження 6.1. Висновки з оберненої граничної теореми. Умови теореми, очевидно, виконано, якщо відомо, що послідовність $\{f_n\left(t\right)\}$ збігається до певної функції $f\left(t\right)$ рівномірно в кожному скінченному інтервалі. Умови теореми, очевидно, виконані також і тоді, коли відомо, що послідовність $\{f_n\left(t\right)\}$ збігається до певної функції $f\left(t\right)$, яка є характеристичною функцією якоїсь випадкової величини.

Приклад 6.12. Як приклад використання граничних теорем доведемо інтегральну теорему Муавра – Лапласа. У прикладі 6.4 знайдено характеристичну функцію випадкової величини

$$\eta = \frac{\mu - np}{\sqrt{npq}}.$$

Вона дорівнює

$$f_n\left(t\right) = \left(qe^{-it\sqrt{\frac{p}{nq}}} + pe^{it\sqrt{\frac{q}{np}}}\right)^n.$$

Скориставшись розкладом функції e^z в ряд Тейлора, знаходимо, що

$$qe^{-it\sqrt{\frac{p}{nq}}}+pe^{it\sqrt{\frac{q}{np}}}=1-\frac{t^2}{2n}\left(1+R_n\right),$$

де

$$R_n=2\sum_{k=3}^{\infty}\frac{1}{k!}\left(\frac{it}{\sqrt{n}}\right)^{k-2}\frac{pq^k+q(-p)^k}{\sqrt{(pq)^k}}.$$

Оскільки $R_n\to 0$ при $n\to\infty$, то маємо

$$f_n\left(t\right)=\left(1-\frac{t^2}{2n}\left(1+R_n\right)\right)^n\to\exp\left(-\frac{t^2}{2}\right).$$

В силу оберненої граничної теореми при всіх значенях x при $n\to\infty$

$$P\left\{\frac{\mu-np}{\sqrt{npq}}<x\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{z^2}{2}}dz.$$

Із нерервності граничної функції випливає, що збіжність рівномірна по x.

Очевидно, що це співвідношення еквівалентно звичайному формулюванню інтегральної теореми Муавра – Лапласа.

6.5 Додатно визначені функції

У попередніх параграфах було встановлено деякі важливі для застосувань властивості характеристичних функцій. У наступному розділі дамо одне з таких застосувань до одного з найбільш розвинених розділів теорії ймовірностей – граничних законів для сум незалежних випадкових величин і разом з тим до теорії однорідних випадкових процесів з незалежними приростами. Мета цього параграфа – вичерпно описати клас характеристичних функцій шляхом доведення теореми, установленої одночасно С. Бохнером і О.Я. Хінчиним, але вперше опублікованої С. Бохнером. Для формулювання та доведення цієї теореми треба ввести нове поняття. Ми скажемо, що функція $f\left(t\right)$ дійсного аргументу t додатно визначена на інтервалі $-\infty<t<+\infty$, якщо б за будь-яких дійсних чисел $t_1,t_2,\ldots,t_n$, комплексних чисел $\xi_1,\xi_2,\ldots,\xi_n$ і цілого числа n виконується таке співвідношення:

$$\sum_{k=1}^{n}\sum_{j=1}^{n}f\left(t_k-t_j\right)\xi_k\overline{\xi}_j\geq 0. \tag{6.22}$$

Перш ніж переходити до викладу основного змісту параграфа, установимо кілька найпростіших властивостей додатно визначених функцій.

1. $f(0) \geq 0$. Справді, покладемо $n = 1, t_1 = 0, \xi_1 = 1$. Тоді з умови додатної визначеності функції знаходимо, що

$$f(t_1 - t_1)\,\xi_1\overline{\xi}_1 = f(0) \geq 0.$$

2. За довільних додатних t

$$f(-t) = \overline{f(t)}.$$

Для доведення цього факту в (6.22) покладемо $n = 2, t_1 = 0, t_2 = t, \xi_1, \xi_2$ – довільні. Маємо:

$$0 \leq \sum_{k=1}^{2}\sum_{j=1}^{2} f(t_k - t_j)\,\xi_k\overline{\xi_j} =$$

$$= f(0-0)\xi_1\overline{\xi_1} + f(0-t)\,\xi_1\overline{\xi_2} + f(t-0)\,\xi_2\overline{\xi_1} + f(t-t)\,\xi_2\overline{\xi_2} =$$

$$= f(0)\left(|\xi_1|^2 + |\xi_2|^2\right) + f(-t)\,\xi_1\overline{\xi_2} + f(t)\,\xi_2\overline{\xi_1.} \qquad (6.23)$$

Тому величина

$$f(-t)\,\xi_1\overline{\xi_2} + f(t)\,\xi_2\overline{\xi_1}$$

має бути дійсною. Таким чином, якщо покласти

$$f(-t) = a_1 + i\beta_1, \quad f(t) = a_2 + i\beta_2,$$

$$\xi_1\overline{\xi_2} = \gamma + i\delta, \quad \xi_2\overline{\xi_1} = \gamma - i\delta,$$

то має бути

$$a_1\delta + \beta_1\gamma - a_2\delta + \beta_2\gamma = 0.$$

Оскільки ξ_1 та ξ_2, а отже, γ та δ довільні, то має бути $a_1 - a_2 = 0$, $\beta_1 + \beta_2 = 0$. Звідси випливає наше твердження.

3. За довільних дійсних t

$$|f(t)| \leq f(0)\,.$$

Покладемо в нерівності (6.23)

$$\xi_1 = -\sqrt{f(t)}, \quad \xi_2 = \sqrt{f(-t)},$$

тоді, згідно з попереднім,

$$f(-t)\,\xi_1\overline{\xi_2}+f(t)\,\xi_2\overline{\xi_1}=-f(t)\,f(-t)-f(t)\,f(-t)=-2|f(t)|^2$$

і

$$|\xi_1|^2=|f(t)|\,,\quad |\xi_2|^2=|f(t)|\,.$$

Нерівність (6.23) набуває вигляду

$$2f(0)\,|f(t)|-2|f(t)|^2\geq 0.$$

Звідси, якщо

$$|f(t)|\neq 0,$$

то

$$f(0)\geq |f(t)|\,.$$

Якщо ж

$$|f(t)|=0,$$

то знову-таки на підставі властивості 1

$$f(0)\geq |f(t)|\,.$$

Звідси випливає, що коли додатно визначена функція така, що $f(0)=0$, то $f(t)=0$.

Теорема 6.15. (Теорема Бохнера – Хінчина.) *Щоб неперервна функція $f(t)$, яка задовольняє умову $f(0)=1$, була характеристичною, необхідно і достатньо, щоб вона була додатно визначеною.*

Доведення. В одному напрямі теорема тривіальна. Справді, якщо $f(t)$ – характеристична функція, то

$$f(t)=\int\limits_{-\infty}^{\infty}e^{itx}dF(x),$$

де $F(x)$ – деяка функція розподілу.

За довільного цілого n і довільних дійсних $t_1,t_2,\ldots,t_n$ та комплексних $\xi_1,\xi_2,\ldots,\xi_n$ маємо

$$\sum_{k=1}^{n}\sum_{j=1}^{n}f(t_k-t_j)\,\xi_k\overline{\xi_j}=\sum_{k=1}^{n}\sum_{j=1}^{n}\int\limits_{-\infty}^{\infty}e^{ix(t_k-t_j)}dF(x)\xi_k\overline{\xi_j}=$$

$$= \int\limits_{-\infty}^{\infty} \sum_{k=1}^{n} \sum_{j=1}^{n} e^{ix(t_k - t_j)} \xi_k \xi_j dF(x) =$$

$$= \int\limits_{-\infty}^{\infty} \left(\sum_{k=1}^{n} e^{it_k x} \xi_k \right) \left(\sum_{j=1}^{n} e^{-it_j x} \xi_j \right) dF(x) =$$

$$= \int\limits_{-\infty}^{\infty} \left| \sum_{k=1}^{n} e^{it_k x} \xi_k \right|^2 dF(x) \geq 0.$$

Доведення достатності потребує складніших міркувань. Розглянемо послідовність чисел

$$c_k = f\left(\frac{k}{n}\right), \quad k = 0, \pm 1, \pm 2, \ldots,$$

залежну від цілого параметра n. У силу додатної визначеності функції $f(x)$ маємо при довільному N

$$P_N^{(n)}(x) = \frac{1}{N} \sum_{k=1}^{N} \sum_{j=1}^{N} f\left(\frac{k-j}{n}\right) e^{-i(k-j)x} \geq 0.$$

Легко підрахувати, що в цій сумі $N - |r|$ доданків, для яких різниця $k - j$ дорівнює r. Далі очевидно, що число $|r|$ може змінюватися від 0 до $N - 1$. Отже, має місце рівність

$$P_N^{(n)}(x) = \sum_{r=-N+1}^{N-1} \left(1 - \frac{|r|}{N}\right) f\left(\frac{r}{n}\right) e^{-irx}.$$

Помножимо обидві частини цієї рівності на e^{isx} і проінтегруємо по x у границях від $-\pi$ до π.

$$\int\limits_{-\pi}^{\pi} e^{isx} P_N^{(n)}(x)\, dx = \sum_{r=-N+1}^{N-1} \left(1 - \frac{|r|}{N}\right) f\left(\frac{r}{n}\right) \int\limits_{-\pi}^{\pi} e^{isx} e^{-irx} dx.$$

Відомо, що

$$\int\limits_{-\pi}^{\pi} e^{-i(r-s)x} dx = \begin{cases} \ \ 0\ , \ r \neq s, \\ 2\pi, \ r = s. \end{cases}$$

Тому

$$\left(1-\frac{|s|}{N}\right) f\left(\frac{r}{n}\right)=\frac{1}{2\pi}\int_{-\pi}^{\pi} P_N^{(n)}\left(x\right) e^{isx}dx=\int_{-\pi}^{\pi} e^{isx}dF_N^{(n)}\left(x\right),$$

де

$$F_N^{(n)}(x)=\frac{1}{2\pi}\int_{-\pi}^{x} P_N^{(n)}\left(z\right) dz$$

є неспадна функція, повна варіація якої дорівнює

$$F_N^{(n)}(\pi)=\frac{1}{2\pi}\int_{-\pi}^{\pi} P_N^{(n)}\left(z\right) dz=f\left(0\right)=1,$$

тобто є функція розподілу.

Нехай тепер $N\to\infty$. Тоді на підставі першої теореми Хеллі можемо знайти послідовність $N_k\to\infty$ при $k\to\infty$, для якої функції $F_{N_k}^{(n)}\left(x\right)$ прямують до граничної функції. Позначимо її через $F^{(n)}\left(x\right)$. Функція $F^{(n)}\left(x\right)$ знову є функцією розподілу, оскільки при довільному N та $\varepsilon>0$ маємо

$$F_{N_k}^{(n)}\left(-\pi-\varepsilon\right)=0,\quad F_{N_k}^{(n)}\left(\pi+\varepsilon\right)=1$$

і, отже, при довільному $\varepsilon>0$ також

$$F^{(n)}\left(-\pi-\varepsilon\right)=0,\quad F^{(n)}\left(\pi+\varepsilon\right)=1.$$

Згідно з другою теоремою Хеллі

$$\lim_{k\to\infty}\int_{-\pi}^{\pi} e^{isx}dF_{N_k}^{(n)}\left(x\right)=\int_{-\pi}^{\pi} e^{isx}dF^{(n)}\left(x\right).$$

Таким чином,

$$f\left(\frac{s}{n}\right)=\int_{-\pi}^{\pi} e^{isx}dF^{(n)}\left(x\right). \tag{6.24}$$

при всіх цілих $s=0,\pm1,\pm2,\ldots$.

Розглянемо тепер послідовність характеристичних функцій $f_n(t)$, визначених рівностями

$$f_n(t) = \int\limits_{-\pi n}^{\pi n} e^{itx} dF_n(x),$$

де

$$F_n(x) = F^{(n)}\left(\frac{x}{n}\right).$$

Легко перевірити, що при всіх цілих k

$$f_n\left(\frac{k}{n}\right) = f\left(\frac{k}{n}\right) \tag{6.25}$$

Але, яке б не було t, можна підібрати таку послідовність $k = k(n, t)$, щоб $0 \le t - \frac{k}{n} < \frac{1}{2\pi}$. Із неперервності функції $f_n(t)$ випливає, що

$$f(t) = \lim_{n\to\infty} f\left(\frac{k}{n}\right) = \lim_{n\to\infty} f_n\left(\frac{k}{n}\right). \tag{6.26}$$

Коли доведемо, що при всіх дійсних t

$$f(t) = \lim_{n\to\infty} f_n(t), \tag{6.27}$$

то доведення теореми буде закінченим, оскільки $f(t)$ неперервна функція і тому в силу оберненої граничної теореми для характеристичних функцій буде характеристичною функцією. Із цією метою зауважимо, що із (6.25) та (6.26) випливає рівність

$$\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty} \left\{\left[f_n(t) - f_n\left(\frac{k}{n}\right)\right] + f_n\left(\frac{k}{n}\right)\right\} =$$

$$= f(t) + \lim_{n\to\infty} \left[f_n(t) - f_n\left(\frac{k}{n}\right)\right]. \tag{6.28}$$

Позначимо $\theta = t - \frac{k}{n}$.

Згідно з вибором величини k маємо $0 \le \theta < \frac{1}{2n}$. За означенням функції $f_n(x)$

$$\left|f_n(t) - f_n\left(\frac{k}{n}\right)\right| = \left|\int\limits_{-\pi n}^{\pi n} e^{i\frac{k}{n}x}\left(e^{i\Theta x} - 1\right) dF_n(x)\right| \le$$

$$\leq \int\limits_{-\pi n}^{\pi n} \left|e^{i\Theta x} - 1\right| dF_n\left(x\right). \tag{6.29}$$

Скориставшись нерівністю Буняковського – Шварца, знаходимо, що

$$\int\limits_{-\pi n}^{\pi n} \left|e^{i\theta x} - 1\right| dF_n\left(x\right) \leq \sqrt{\int\limits_{-\pi n}^{\pi n} \left|e^{i\theta x} - 1\right|^2 dF_n\left(x\right)} =$$

$$= \sqrt{\int\limits_{-\pi n}^{\pi n} 2\left(1 - \cos\theta x\right) dF_n\left(x\right)} = \sqrt{2\left(1 - Rf_n\left(\theta\right)\right)}, \tag{6.30}$$

де символ $Rf_n\left(\theta\right)$ означає дійсну частину $f_n\left(\theta\right)$. Так як $\cos z \leq \cos az$ при $0 \leq a \leq \frac{1}{2}$ та $-\pi \leq z \leq \pi$, то

$$\begin{array}{c} 1 - Rf_n\left(\theta\right) = \int\limits_{-\pi n}^{\pi n} \left(1 - \cos\theta x\right) dF_n\left(x\right) = \int\limits_{-\pi}^{\pi} \left(1 - \cos n\theta z\right) dF_n\left(zn\right) \leq \\ \leq \int\limits_{-\pi}^{\pi} (1 - \cos z) dF_n\left(zn\right) \end{array}.$$

Оскільки

$$F_n\left(zn\right) = F^{(n)}\left(z\right),$$

то

$$1 - Rf_n\left(\Theta\right) = \int\limits_{-\pi}^{\pi} \left(1 - \cos z\right) dF^{(n)}\left(z\right) = 1 - R\int\limits_{-\pi}^{\pi} e^{iz} dF^{(n)}\left(z\right).$$

Звідси на підставі (6.25) знаходимо, що

$$1 - Rf_n\left(\theta\right) \leq 1 - Rf\left(\frac{1}{n}\right).$$

Зіставивши цю нерівність з нерівностями (6.28) і (6.29), знаходимо, що

$$\left|f_n\left(t\right) - f_n\left(\frac{k}{n}\right)\right| \leq \sqrt{2\left(1 - Rf\left(\frac{1}{n}\right)\right)}.$$

Через неперервність функції

$$f\left(t\right)$$

із попередньої нерівності випливає, що

$$\lim_{n\to\infty}\left[f_n\left(t\right)-f_n\left(\frac{k}{n}\right)\right]=0.$$

Співвідношення (6.27), як це видно тепер із (6.28), доведено. $\square$

6.6 Характеристичні функції багатовимірних випадкових величин

У цьому параграфі подаємо без доведення основні властивості характеристичних функцій багатовимірних випадкових величин. Характеристичною функцією n-вимірної випадкової величини

$$(\xi_1,\xi_2,\ldots,\xi_n)$$

називається математичне сподівання величини

$$e^{i(t_1\xi_1+t_2\xi_2+\ldots t_n\xi_n)},$$

де $t_1,t_2,\ldots,t_n$ – дійсні змінні:

$$f\left(t_1,t_2,\ldots,t_n\right)=\mathsf{E}e^{i\sum\limits_{k=1}^{n}\xi_k t_k}. \tag{6.31}$$

Якщо $F\left(x_1,x_2,\ldots,x_n\right)$ є функція розподілу величини $(\xi_1,\xi_2,\ldots,\xi_n)$, то, як видно із зазначеного вище,

$$f\left(t_1,t_2,\ldots,t_n\right)=\int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty}e^{i\sum\limits_{k=1}^{n}x_k t_k}dF\left(x_1,x_2,\ldots,x_n\right). \tag{6.32}$$

Подібно до того, як це було доведено в одновимірному випадку, можна показати, що характеристична функція n-вимірної випадкової величини рівномірно неперервна у всьому просторі $(-\infty<t_j<\infty,1\le j\le n)$ і задовольняє такі співвідношення:

$$\left|f\left(0,0,\ldots 0,\right)\right|=1,$$

$$\left|f\left(t_1,t_2,\ldots,t_n\right)\right|\le 1,\quad -\infty<t_j<\infty,1\le j\le n,$$

$$f\left(-t_1,-t_2,\ldots,-t_n\right)=\overline{f\left(t_1,t_2,\ldots,t_n\right)}.$$

За характеристичною функцією $f(t_1, t_2, \ldots, t_n)$ випадкової величини $(\xi_1, \xi_2, \ldots, \xi_n)$ легко знайти характеристичну функцію k-вимірної ($k < n$) величини $(\xi_{j_1}, \xi_{j_2}, \ldots, \xi_{j_k})$, компонентами якої є величини ξ_s $(1 \le s \le n)$. Для цього у формулі (6.31) треба покласти рівними нулю всі аргументи ξ_s за $s \neq j_r$ $(1 \le r \le k)$. Так, наприклад, характеристична функція величини ξ_1 дорівнює

$$f(t_1) = f(t_1, 0, \ldots, 0).$$

З означення випливає, що коли компоненти величини $(\xi_1, \xi_2, \ldots, \xi_n)$ є незалежними випадковими величинами, то її характеристична функція дорівнює добуткові характеристичних функцій її компонент

$$f(t_1, t_2, \ldots, t_n) = f(t_1)\, f(t_2) \cdots f(t_n).$$

Так само як і в одновимірному випадку, багатовимірні характеристичні функції дозволяють легко знаходити моменти різних порядків. Так, наприклад,

$$\mathsf{E}\xi_1^{k_1}\xi_2^{k_2}\ldots\xi_n^{k_n} =$$
$$= (i)^{-\sum\limits_{j=1}^{n} k_j} \left[\frac{\partial^{k_1+k_2+\ldots k_n} f(t_1, t_2, \ldots t_n)}{(\partial t_1)^{k_1}(\partial t_2)^{k_2}\ldots(\partial t_n)^{k_n}}\right]_{t_1=t_2=\ldots t_n=0}.$$

Якщо позначити

$$\psi(t_1, t_2, \ldots, t_n) = \ln f(t_1, t_2, \ldots, t_n),$$

то легко підрахувати, що

$$M\xi_j = -i\left[\frac{\partial\psi}{\partial t_j}\right]_{t_1=t_2=\ldots=t_n=0},$$

$$\operatorname{cov}(\xi_j, \xi_k) = -\left[\frac{\partial^2\psi}{\partial t_j \partial t_k}\right]_{t_1=t_2=\ldots=t_n=0}.$$

Для обчислення характеристичних функцій корисно знати таку теорему, яку читач легко може довести самостійно.

Теорема 6.16. *Якщо характеристична функція величини* $(\xi_1, \xi_2, \ldots, \xi_n)$ *дорівнює* $f(t_1, t_2, \ldots, t_n)$, *то характеристична функція величини*

$$(a_1\xi_1 + b_1, a_2\xi_2 + b_2, \ldots, a_n\xi_n + b_n),$$

де a_j *та* b_j $(1 \le j \le n)$ *– дійсні сталі, дорівнює*

$$e^{i\sum\limits_{k=1}^{n} b_k t_k} f\left(a_1 t_1, a_2 t_2, \ldots, a_n t_n\right).$$

Приклад 6.13. Обчислимо характеристичну функцію двовимірної випадкової величини, розподіленої за нормальним законом

$$p\left(x, y\right) = \frac{1}{2\pi\sqrt{1-r^2}} e^{-\frac{1}{2\left(1-r^2\right)}\left[x^2-2rxy+y^2\right]} \tag{6.33}$$

Замінами

$$u = \frac{y - rx - it_2\left(1-r^2\right)}{\sqrt{1-r^2}}, \quad v = x - i\left(t_1 + rt_2\right)$$

зведемо $f\left(t_1, t_2\right)$ до вигляду

$$f\left(t_1, t_2\right) = e^{-\frac{1}{2}\left(t_1^2+2rt_1t_2+t_2^2\right)} \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\left(u^2+v^2\right)} du dv =$$

$$= e^{-\frac{1}{2}\left(t_1^2+2rt_1t_2+t_2^2\right)}.$$

Приклад 6.14. Знайдемо характеристичну функцію величини (η_1, η_2), розподіленої за нормальним законом:

$$\pi\left(x, y\right) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2\left(1-r^2\right)}\left[\frac{(x-a)^2}{\sigma_1^2} - 2r\frac{(x-a)(x-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{\sigma_2^2}\right]}. \tag{6.34}$$

Якщо покладемо

$$\eta_1 = \sigma_1\xi_1 + a, \quad \eta_2 = \sigma_2\xi_2 + b,$$

то величина (ξ_1, ξ_2) буде розподілена за законом (6.33). Згідно з теоремою 1 характеристична функція величини (η_1, η_2) дорівнює

$$\varphi\left(t_1, t_2\right) = e^{iat_1 + ibt_2 - \frac{1}{2}\left(\sigma_1^2 t_1^2 + 2r\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2\right)}.$$

З означення характеристичної функції випливає таке твердження:

Теорема 6.17. *Якщо $f(t_1,\dots,t_n)$ є характеристичною функцією величини $(\xi_1,\xi_2,\dots,\xi_n)$, то характеристична функція суми $(\xi_1+\xi_2+\dots+\xi_n)$ дорівнює*

$$f(t) = f(t,t,\dots,t).$$

Зауваження 6.2. Зауважимо, що

$$f_1(t) = f(tt_1, tt_2,\dots,tt_n)$$

є характеристична функція суми

$$t_1\xi_1 + t_2\xi_2 + \dots + t_n\xi_n.$$

Приклад 6.15. Застосуємо теорему 6.17 до визначення функції розподілу суми $\eta_1+\eta_2$, якщо (η_1,η_2) розподілені за законом (6.34). Згідно з теоремою 2 характеристична функція суми $\eta_1+\eta_2$ дорівнює

$$f(t) = e^{it(a+b)-\frac{1}{2}t^2\left(\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2\right)}.$$

Це характеристична функція нормального закону з математичним сподіванням $a+b$ та дисперсією

$$\sigma_1^2+2r\sigma_1\sigma_2+\sigma_2^2.$$

Цей результат було добуто раніше безпосередньо.

Теорема 6.18. *Функція розподілу $F(x_1,x_2,\dots,x_n)$ однозначно визначаєтьcя своєю характеристичною функцією.*

Доведення цього твердження спирається на формулу обернення.

Теорема 6.19. (Формула обернення). *Якщо $f(t_1,t_2,\dots,t_n)$ – характеристична функція, а $F(x_1,x_2,\dots,x_n)$ – функція розподілу випадкової величини $(\xi_1,\xi_2,\dots,\xi_n)$, то*

$$P\{a_k \le \xi_k < b_k; k=1,2,\dots,n\} =$$
$$= \lim_{T\to\infty} \frac{1}{(2\pi)^n} \int_{-T}^{T} \dots \int_{-T}^{T} \prod_{k=1}^{n} \frac{e^{-ita_k}-e^{-itb_k}}{it_k} f(t_1,\dots,t_n)\,dt_1\dots dt_n,$$

де a_k та b_k– довільні дійсні числа, що задовольняють єдину вимогу: імовірність потрапляння на поверхню паралелепіпеда

$$a_k \le \xi_k \le b_k, \quad 1 \le k \le n,$$

дорівнює нулю.

Подібно до того, як це було зроблено в одновимірному випадку, можна довести дуже важливу теорему.

Теорема 6.20. *Щоб послідовність* $\{F_k(x_1, x_2, \ldots, x_n)\}$ *функцій розподілу при* $k \to \infty$ *збігалася до* n*-вимірної функції розподілу* $F(x_1, x_2, \ldots, x_n)$*, необхідно й достатньо, щоб відповідні характеристичні функції* $\{f_k(t_1, t_2, \ldots, t_n)\}$ *при* $k \to \infty$ *рівномірно в кожному скінченному паралелепіпеді* $|t_j| \le T_j$, $1 \le j \le n$*, збігалися до граничної функції* $f(t_1, t_2, \ldots, t_n)$*. Ця функція є характеристичною функцією для граничної функції розподілу.*

7 Класична гранична теорема

Інтегральна гранична теорема Муавра – Лапласа для послідовності незадежних дослідів Бернуллі стала джерелом великого циклу досліджень, які мають фундаментальне з начення як для самої теорії ймовірностей, так і для її численних застосувань у природничих, економічних і технічних науках. Щоб скласти уявлення про напрям цих досліджень, надамо запису теореми Муавра – Лапласа трохи іншого вигляду. А саме, якщо, як уже неодноразово було зроблено, через μ_k позначити кількість появ події A в k-му випробовуванні, то кількість появ події A в n послідовних випробуваннях дорівнює

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

Математичне сподівання та дисперсія цієї випадкової велечини дорівнюють

$$\sum_{k=1}^{n} \mathsf{E}\mu_k = np, \quad \sum_{k=1}^{n} D\mu_k = npq.$$

Враховуючи такі значення математичного сподівання та дисперсії, теорема Муавра – Лапласа може бути записана в такому вигляді.

Теорема 7.1. (Теорема Муавра – Лапласа.) *При* $n \to \infty$

$$P\left\{ a \le \frac{\sum\limits_{k=1}^{n} (\mu_k - \mathsf{E}\mu_k)}{\sqrt{\sum\limits_{k=1}^{n} D\mu_k}} < b \right\} \to \frac{1}{\sqrt{2\pi}} \int\limits_{a}^{b} e^{-\frac{z^2}{2}} dz. \tag{7.1}$$

Іншими словами: Імовірність того, що сума відхилень значень незалежних випадкових величин, які набувають двох значень 0 та 1 з імовірностями p та q, $(q = 1 - p, 0 < p < 1)$, *від математичних сподівань цих величин, поділена на корінь квадратний із суми дисперсій доданків, буде міститися в межах від a до b, зі збільшенням кількості доданків до нескінченності рівномірно відносно a та b збігається до інтеграла*

$$\frac{1}{\sqrt{2\pi}} \int\limits_{a}^{b} e^{-\frac{z^2}{2}} dz.$$

Природно виникають запитання: як тісно пов'язано співвідношення (7.1) зі спеціальним вибором доданків, і чи не буде воно справджуватись і при слабших обмеженнях на функції розподілу доданків. Постановка цієї задачі, а також її розв'язання значною мірою викладена в працях П.Л.Чебишева та його учнів – А.А.Маркова і О.М.Ляпунова. Унаслідок їхніх досліджень виявилося, що на доданки слід накласти лише найзагальніші обмеження, зміст яких полягає в тому, що коли окремі доданки мають малий вплив на суму, то неминуче виконується співвідношення (7.1).

У наступному розділі дано доведення Ліндеберга теореми Ляпунова. Не будемо наголошувати на математичному інтересі до цих досліджень, адже загальність і закінченість їх говорять самі за себе. Значно важливіше, на наш погляд, виявити причини, завдяки яким отримані результати набули великого значення в застосуваннях. Одна з найважливіших схем, за якою йде використання результатів теорій імовірностей в природознавстві й техніці, полягає ось у чому: вважають, що процес проходить під впливом великої кількості незалежно діючих випадкових факторів, кожний з яких мало змінює хід явища або процесу. Дослідник, який цікавиться вивченням процесу в цілому, а не дією окремих факторів, спостерігає лише сумарну дію цих факторів.

Подамо два типових приклади.

Приклад 7.1. Коли проводиться якесь вимірювання, то на його результат впливає велика кількість факторів, які неможливо всі врахувати і які породжують похибки у вимірах. До них належать похибки, викликані станом приладу, який може дещо змінюватися під впливом різних атмосферних змін або механічних причин, а також особисті помилки дослідника, зумовлені особливостями його зору або слуху, які також можуть дещо змінюватися залежно від його психічного або фізичного стану. Кожний із цих факторів, зокрема, викликав би незначну "елементарну похибку". Але на результат виміру впливають разом усі ці "елементарні похибки" в сукупності. Інакше кажучи, похибка, яку фактично спостерігають, буде випадковою величиною, яка є сумою величезної кількості незначних за величиною і незалежних між собою випадкових величин. І хоч ці випадкові величини невідомі, так само як невідомі їхні функції розподілу, їхній вплив на результат виміру значний і тому має підлягати вивченню.

Приклад 7.2. У процесі масового виробництва в багатьох галузях промисловості виробляється велика кількість однакових предметів. Нас цікавить певна числова характеристика виробу. За певної величини вибраної нами характеристики виріб відповідає технічним нормам. Проте у практиці завжди спостерігається деяке відхилення від цієї величини. За правильно поставленого процесу виробництва такі відхилення можуть викликатися лише випадковими причинами, причому ефект дії кожної з них незначний. Але сумарна дія спричиняє помітне відхилення від норми.

Таких прикладів можна навести скільки завгодно: положення, швидкість і напрямок руху молекули газу, які визначаються великою кількістю зіткнень з іншими молекулами; розсіювання точок влучання снарядів при стрільбі з гармати, викликане величезною кількістю причин: коливанням у кількостях вибухової речовини, неправильністю форми стінок снаряда, коливанням атмосферних умов, неточністю в наводці та ін. Взагалі випадкові відхилення, викликані великою кількістю незалежних між собою незначних випадкових відхилень, зустрічаються при вивченні різноманітних явищ природи і технічних процесів. Таким чином, виникає необхідність вивчення закономірностей, властивих сумам великої кількості незалежних випадкових величин, кожна з яких мало впливає на суму. Цій останній вимозі надамо пізніше точнішого змісту. Замість того, щоб вивчати суму великої, але скінченної кількості доданків, розглядатимемо послідовність сум з дедалі більшою кількістю доданків і вважатимемо, що розв'язання задач, які нас цікавлять, даються граничними функціями для функцій розподілу сум. Такий перехід від скінченної постановки задачі до граничної є звичайним як для сучасної математики, так і для багатьох розділів природознавства.

Отже, ми прийшли до такої задачі.

Дана послідовність

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

взаємно незалежних випадкових величин, щодо яких припускаємо, що вони мають скінченні математичні сподівання та дисперсії. Далі будемо дотримуватися таких позначень:

$$a_k = \mathsf{E}\xi_k, \quad b_k^2 = D\xi_k, \quad B_n^2 = \sum_{k=1}^{n} b_k^2 = D\left(\sum_{k=1}^{n} \xi_k\right).$$

Постає питання, які умови треба накласти на величини ξ_k, щоб функції розподілу величин

$$\frac{1}{B_n}\sum_{k=1}^{n}\left(\xi_k - a_k\right) \tag{7.2}$$

при $n \to \infty$ збігалися до нормального закону розподілу.

У наступному параграфі побачимо, що для цього достатньо, щоб виконувалася умова Ліндеберга: для довільного $\tau > 0$

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{B_n^2}\int\limits_{|x-a_k|>\tau B_n}\left(x-a_k\right)^2 dF_k\left(x\right) = 0,$$

де $dF_k\left(x\right)$ означає функцію розподілу величини ξ_k. Наша найближча мета полягає в тому, щоб виявити сенс цієї умови і показати, що вона замінює вимогу рівномірної малості доданків

$$\frac{1}{B_n}\left(\xi_k - a_k\right)$$

в сумі (7.2). Справді, нехай $\tau > 0$ – довільне додатне число. Знайдемо ймовірність того, що хоч би одна величина

$$\frac{1}{B_n}\left(\xi_k - a_k\right), \quad k = 1, 2, \ldots n,$$

перевищить за абсолютною величиною τ. Маємо

$$P\left\{\max_{1\le k\le n}\left|\xi_k - a_k\right| > \tau B_n\right\} = P\left\{A_1 \cup A_2 \cup \cdots \cup A_n\right\},$$

де $A_k,\ k = 1, 2, \ldots, n$, означає подію, яка полягає в тому, що

$$\left|\xi_k - a_k\right| > \tau B_n.$$

Із попередніх розділів відомо, що

$$P\left\{A_1 \cup A_2 \cup \cdots \cup A_n\right\} \le \sum_{k=1}^{n} P\left\{A_k\right\}.$$

Але

$$P\left(A_k\right) = \int\limits_{|x-a_k|>\tau B_n} dF_k\left(x\right) \le \frac{1}{\left(\tau B_n\right)^2}\int\limits_{|x-a_k|>\tau B_n}\left(x-a_k\right)^2 dF_k\left(x\right),$$

тому

$$P\left\{\max_{1\le k\le n}|\xi_k-a_k|>\tau B_n\right\}\le\frac{1}{\tau^2B_n^2}\sum_{k=1}^{n}\int\limits_{|x-a_k|>\tau B_n}(x-a_k)^2dF_k(x).$$

За умови Ліндеберга, яка б не була стала $\tau>0$, остання сума при $n\to\infty$ прямує до нуля. Цим і доведено те, що було потрібно.

Зауваження 7.1. Зауважимо, що виявлений нами сенс умов, які слід накласти на доданки, щоб функції розподілу сум (7.2) збігалися до нормального розподілу, був повною мірою з'ясований в дослідженнях А. А. Маркова та О. М. Ляпунова.

7.1 Теорема Ліндеберга

Теорема 7.2. (Теорема Ліндеберга.) *Якщо послідовність взаємно незалежних випадкових величин*

$$\xi_1,\xi_2,\dots,\xi_n,\dots$$

при довільному сталому $\tau>0$ задовольняє умову Ліндеберга:

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int\limits_{|x-a_k|>\tau B_n}(x-a_k)^2dF_k(x)=0,\tag{7.3}$$

де

$$a_k=\mathsf{E}\xi_k,\quad B_n^2=\sum_{k=1}^{n}D\xi_k,\quad F_k(x)=P\left\{\xi_k<x\right\},$$

то рівномірно відносно x при $n\to\infty$ має місце співвідношення

$$P\left\{\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k-a_k)<x\right\}\to\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-\frac{z^2}{2}}dz.\tag{7.4}$$

Доведення. Для скорочення запису введемо позначення

$$\xi_{nk}=\frac{\xi_k-a_k}{B_n},\quad F_{nk}(x)=P\left\{\xi_{nk}<x\right\}.$$

Очевидно, що

$$\mathsf{E}\xi_{nk}=0,\quad D\xi_{nk}=\frac{D\xi_k}{B_n^2},$$

і, отже,

$$\sum_{k=1}^{n} D\xi_{nk} = 1.$$

У цих позначеннях умова Ліндеберга набуває такого вигляду:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int\limits_{|x|>\tau} x^2 dF_{nk}(x) = 0. \tag{7.5}$$

Характеристична функція суми

$$\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k) = \sum_{k=1}^{n}\xi_{nk}$$

дорівнює

$$\varphi_n(t) = \prod_{k=1}^{n} f_{nk}(t).$$

Треба довести, що

$$\lim_{n\to\infty}\varphi_n(t) = e^{-\frac{t^2}{2}}.$$

Для цього покажемо насамперед, що множники $f_{nk}(t)$ при $n \to \infty$ рівномірно відносно k, $1 \le k \le n$, прямують до одиниці. Справді, взявши до уваги рівність

$$\mathsf{E}\xi_{nk} = 0,$$

знайдемо, що

$$f_{nk}(t) - 1 = \int\limits_{-\infty}^{\infty} \left(e^{itx} - 1 - itx\right) dF_{nk}(x).$$

Із того що при довільному дійсному a

$$\left|e^{ia} - 1 - ia\right| \le \frac{a^2}{2} \tag{7.6}$$

маємо нерівність

$$|f_{nk}(t) - 1| \le \frac{t^2}{2}\int\limits_{-\infty}^{\infty} x^2 dF_{nk}(x).$$

Цю та подібні нерівності можна вивести хоч би й таким шляхом. Із того, що

$$\left|e^{ia}-1\right| = \left|\int\limits_0^a e^{ix}dx\right| \le |a|$$

маємо нерівність

$$\left|e^{ia}-1-ia\right| = \left|\int\limits_0^a \left(e^{ix}-1\right)dx\right| \le \frac{a^2}{2}.$$

З останньої нерівності випливає

$$\left|e^{ia}-1-ia+\frac{a^2}{2}\right| = \left|\int\limits_o^a \left(e^{ix}-1-ix\right)dx\right| \le$$

$$\le \int\limits_0^{|a|} \left|e^{ix}-1-ix\right|dx \le \int\limits_0^{|a|} \frac{1}{2}x^2dx = \frac{1}{6}|a|^3. \tag{7.7}$$

Нехай ε – довільне додатне число. Тоді очевидно, що

$$\int\limits_{-\infty}^{\infty} x^2dF_{nk}(x) = \int\limits_{|x|\le\varepsilon} x^2dF_{nk}(x) + \int\limits_{|x|>\varepsilon} x^2dF_{nk}(x)$$

$$\le \varepsilon^2 + \int\limits_{|x|>\varepsilon} x^2dF_{nk}(x).$$

Останній доданок згідно із (7.5) за досить великих n можна зробити меншим від ε^2. Таким чином, для всіх досить великих n рівномірно відносно k, $1 \le k \le n$, у довільному скінченному інтервалі $|t| \le T$

$$|f_{nk}(t)-1| < \varepsilon^2T^2.$$

Звідси виводимо, що рівномірно відносно k, $1 \le k \le n$,

$$\lim_{n\to\infty} f_{nk}(t) = 1, \tag{7.8}$$

і що для всіх досить великих n і t у довільному скінченному інтервалі $|t| \le T$

$$|f_{nk}(t)-1| < \frac{1}{2}. \tag{7.9}$$

Отже, можна в інтервалі $|t| \le T$ записати, що

$$\ln \varphi_n(t) = \sum_{k=1}^{n} \ln f_{nk}(t) = \sum_{k=1}^{n} \ln\left[1 + \left(f_{nk}(t) - 1\right)\right] =$$

$$= \sum_{k=1}^{n} \left(f_{nk}(t) - 1\right) + R_n, \tag{7.10}$$

де

$$R_n = \sum_{k=1}^{n} \sum_{s=2}^{\infty} \tfrac{(-1)^s}{s} \left(f_{nk}(t) - 1\right)^s.$$

Відповідно до (7.9)

$$|R_n| \le \sum_{k=1}^{n} \sum_{s=2}^{\infty} \frac{1}{2} |f_{nk}(t) - 1|^s = \frac{1}{2} \sum_{k=1}^{n} \frac{|f_{nk}(t) - 1|^2}{1 - |f_{nk}(t) - 1|} \le$$

$$\le \sum_{k=1}^{n} |f_{nk}(t) - 1|^2,$$

Оскільки

$$\sum_{k=1}^{n} |f_{nk}(t) - 1| = \sum_{k=1}^{n} \left| \int_{-\infty}^{\infty} \left(e^{itx} - 1 - itx\right) dF_{nk}(x) \right| \le$$

$$\le \frac{t^2}{2} \sum_{k=1}^{n} \int_{-\infty}^{\infty} x^2 dF_{nk}(x) = \frac{t^2}{2},$$

то

$$|R_n| \le \frac{t^2}{2} \max_{1 \le k \le n} |f_{nk}(t) - 1|.$$

Із (7.8) випливає, що рівномірно відносно t в кожному скінченному інтервалі $|t| \le T$ при $n \to \infty$ маємо

$$R_n \to 0. \tag{7.11}$$

Можна записати

$$\sum_{k=1}^{n} \left(f_{nk}(t) - 1\right) = -\frac{t^2}{2} + \rho_n, \tag{7.12}$$

де

$$\rho_n = \sum_{k=1}^{n} \int\limits_{-\infty}^{\infty} \left(e^{itx} - 1 - itx - \frac{(itx)^2}{2} \right) dF_{nk}(x).$$

Нерівності (7.6) та (7.7) дозволяють отримати таку оцінку

$$|\rho_n| \le \sum_{k=1}^{n} \int\limits_{|x|\le\varepsilon} \left| e^{itx} - 1 - itx - \frac{(itx)^2}{2} \right| dF_{nk}(x) +$$

$$+ \sum_{k=1}^{n} \int\limits_{|x|>\varepsilon} \left| e^{itx} - 1 - itx - \frac{(itx)^2}{2} \right| dF_{nk}(x) \le$$

$$\le \frac{|t|^3}{6} \sum_{k=1}^{n} \int\limits_{|x|\le\varepsilon} |x|^3 dF_{nk}(x) + t^2 \sum_{k=1}^{n} \int\limits_{|x|>\varepsilon} x^2 dF_{nk}(x) \le$$

$$\le \frac{|t|^3}{6} \cdot \varepsilon + t^2 \left(1 - \frac{|t|}{6} \cdot \varepsilon \right) \sum_{k=1}^{n} \int\limits_{|x|>\varepsilon} x^2 dF_{nk}(x).$$

Згідно з умовою (7.5) другий доданок при довільному $\varepsilon > 0$ може бути зроблений меншим від довільного $\eta > 0$, якщо тільки n досить велике, а так як $\varepsilon > 0$ довільне, то можна його вибрати таким малим, що яке б не було $\eta > 0$ і T, для всіх t, які містяться в інтервалі $|t| \le T$, виконувалася б нерівність

$$|\rho_n| < 2\eta.$$

Ця нерівність показує, що рівномірно в кожному скінченному інтервалі

$$\lim_{n\to\infty} \rho_n = 0. \tag{7.13}$$

Зіставивши співвідношення (7.10), (7.11), (7.12), (7.13) отримуємо, що рівномірно в кожному скінченному інтервалі

$$\lim_{n\to\infty} \ln \varphi_n(t) = -\frac{t^2}{2}.$$

Цим, очевидно, завершується доведення теореми. □

Як наслідок із доведеної теореми дістаємо таку теорему.

Теорема 7.3. *Якщо незалежні випадкові величини*

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

однаково розподілені та мають скінченні математичні сподівання $\mathsf{E}\xi_k = a$ *та скінченні ненульові дисперсії* $D\xi_k = b^2$*, то при* $n \to \infty$ *рівномірно по* x

$$P\left\{\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a) < x\right\} \to \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{z^2}{2}}dz.$$

Доведення. За умов теореми маємо

$$B_n^2 = \sum_{k=1}^{n} b_k^2 = nb^2.$$

Умова Ліндеберга має вигляв

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int\limits_{|x-a|>\tau B_n}|x-a|^2 dF_k(x) =$$

$$= \frac{1}{nb^2}n\int\limits_{|x-a|>\tau B_n}|x-a|^2 dF_1(x) =$$

$$= \frac{1}{b^2}\int\limits_{|x-a|>\tau b\sqrt{n}}|x-a|^2 dF_1(x).$$

Останній вираз прямує до нуля при $n \to \infty$ оскільки за умов теореми випадкові величини мають скінченні ненульові дисперсії $D\xi_k = b^2$, $0 < b < \infty$. □

Вказану теорему можна довести спираючись на обернену граничну терему для характерисимчних функцій.

Доведення. Дійсно, виначимо випадкові величини

$$\zeta_n = \frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a) = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\frac{\xi_k - a}{b} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\eta_k,$$

де

$$\eta_k = \frac{\xi_k - a}{b}, \quad \mathsf{E}\eta_k = 0, \quad \mathsf{D}\eta_k = 1.$$

Згідно з теоремою 6.6 харакреристична функція величини $\frac{\eta_k}{\sqrt{n}}$ при $n \to \infty$ може бути записана у вигляді

$$f_{\frac{\eta_k}{\sqrt{n}}}(t) = 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{2n}\right).$$

Оскільки

$$f_{\zeta_k}(t) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{2n}\right)\right)^n \to \exp\left(-\frac{t^2}{2}\right).$$

В силу оберненої граничної теореми при всіх значенях x при $n \to \infty$

$$P\left\{\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a) < x\right\} \to \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{z^2}{2}}dz.$$

Із нерервності граничної функції випливає, що збіжність рівномірна по x.

$\square$

7.2 Теорема Ляпунова

Як наслідок із теореми Ліндеберга дістаємо теорему Ляпунова.

Теорема 7.4. (Теорема Ляпунова.) *Якщо для послідовності незалежних випадкових величин*

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

можна підібрати таке додатне число $\delta > 0$, що при $n \to \infty$

$$\frac{1}{B_n^{2+\delta}}\sum_{k=1}^{n} E|\xi_k - a_k|^{2+\delta} \to 0, \tag{7.14}$$

то яке б не було x при $n \to \infty$

$$P\left\{\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k) < x\right\} \to \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{z^2}{2}}dz.$$

Доведення. При довільному $r > 0$ мають місце нерівності

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int\limits_{|x-a_k|>\tau B_n}|x-a_k|^2 dF_k(x) \le$$

$$\le \frac{1}{B_n^2(rB_n)^\delta}\sum_{k=1}^{n}\int\limits_{|x-a_k|>\tau B_n}|x-a_k|^{2+\delta} dF_k(x) \le$$

$$\le \frac{1}{\tau^\delta}\cdot\frac{1}{B_n^{2+\delta}}\sum_{k=1}^{n}\int\limits_{-\infty}^{\infty}|x-a_k|^{2+\delta} dF_k(x).$$

Отже, із умови Ляпунова (7.14) випливає умова Ліндеберга (7.3). Цим, очевидно, теорема Ляпунова доведена.

$\square$