МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

П. П. КостробІй

ТЕОРІЯ ЙМОВІРНОСТЕЙ

Львів
Видавництво
Національного університету «Львівська політехніка»
2015

УДК 519.21

ЛЕКЦІЯ 1

Випадкові події. Ймовірності випадкових подій

§1. Історичний вступ

У своїй практичній діяльності людина на кожному кроці зустрічається з явищами, які надалі будемо називати *випадковими*. Найпростішим прикладом випадкових явищ є помилки вимірювань. Проводячи вимірювання однієї і тієї ж величини ми завжди отримуємо близькі, але різні результати. Це пояснюється тим, що результат кожного вимірювання містить похибку. Передбачити, якою буде похибка даного конкретного вимірювання, або навіть визначити її після вимірювання, принципово неможливо.

Проводячи експериментальні дослідження будь-якого явища і систематизуючи результати експерименту в вигляді графічних залежностей, ми переконуємось, що експериментальні точки (якщо їх достатньо багато) ніколи не лягають на одну криву, а завжди заповнюють певну смугу, тобто існує випадковий розкид експериментальних точок. Цей розкид пояснюється як помилкою вимірювань, так і дією інших випадковостей.

При експериментальному дослідженні будь-якого явища з метою встановлення його закономірностей слід спостерігати це явище багатократно за *однакових умовах.* Під *однаковими умовами* ми будемо розуміти одинакові значення всіх кількісних характеристик контрольованих факторів. Всі неконтрольовані фактори при цьому будуть різними, при чому передбачити заздалегідь, якими вони будуть при даному конкретному вимірюванні, принципово неможливо.

Однак при багатократному спостереженні випадкових явищ в них самих, можна зауважити закономірності. Вивчаючи ці закономірності, ми можемо в якісь мірі керувати ними, обмежуючи їх вплив і результати їх дій, та навіть, якісь з них цілеспрямовано використовувати.

Вивченням закономірностей випадкових явищ, що піддаються багатократному спостереженню займається особливий розділ математики — *теорія ймовірностей*, яка не поступається іншим математичним дисциплінам. Бона грунтується на своїй системі основних понять та аксіом, з яких строго виводяться всі її результати. Започаткована у середині XVII століття такими вченими як Гюйгенс (1629–1695), Паскаль (1623–1662), Ферма (1601–1665), Якоб Бернуллі (1654–1705), теорія ймовірностей зазнала справжнього розквіту у XX столітті. Знайдено її тісні зв'язки з іншими розділами математики, широкий спектр практичних застосувань. Сьогодні ймовірнісні методи стали повсякденними не тільки в математиці, фізиці чи економіці, а й у медицині, лінгвістиці, археології. Математична обробка експериментальних даних, теорія планування експерименту та багато інших прикладних наук базуються на теорії ймовірностей.

До 20-х років минулого століття загальноприйнятою була думка, що теорія ймовірностей застосовується для опису тільки тих явищ, відносно яких ми не володіємо повною інформацією. Вважалося, що врахувавши усі необхідні фактори, ми завжди прийдемо до строго детермінованого опису з однозначним передбаченням результатів. Проте створення квантової механіки привело до відкриття фундаментального факту: адекватний опис фізики мікросвіту можливий тільки на основі ймовірнісної концепції, — і це не пов'язано з неповнотою наших знань, а лежить у самій природі речей. За образним висловом А. Ейнштейна, виявилося, що сам Бог грає в кості!

І справді, чи ж не дивно, що мова, яка виникла і початково розвивалася з потреб таких азартних ігор, як гра в кості та карти, виявилася єдиною мовою, придатною для квантової фізики — фундаменту сучасного природознавства? Воістину, пророчі слова написав Гюйгенс у 1657 році у своєму трактаті «Про підрахунки в азартній грі»: «Я думаю, що уважно вивчаючи предмет, читач помітить, що має справу не тільки з грою, а що тут закладаються основи дуже цікавої і глибокої теорії».

Наостанок зазначимо, що великий внесок у розвиток теорії ймовірностей зробили російські та українські вчені. Перший в Росії курс теорії ймовірностей, який справив значний вплив на розвиток інтересу до цієї науки, був написаний В. Я. Буняковським (1804–1889). Вчений широко культивував застосування теорії ймовірностей у страховій справі і демографії. Його діяльність підготувала передумови для виникнення і розвитку російської ймовірнісної школи, пов'язаної з іменами П. Л. Чебишева (1821–1894), А. А. Маркова (1856–1922), А. М. Ляпунова (1857–1918). Визначні результати в теорії ймовірностей уже в наш час одержали вчені С. Н. Бернштейн (1880–1968), А. М. Колмогоров (1903–1987), А. Я. Хінчин (1894–1959), В. І. Гливенко (1897–1940), А. В. Скороход (1930–2011) та інші.

§2. Ймовірнісні експерименти

Підкинемо декілька раз монету, фіксуючи щоразу після падіння її верхню сторону. Це може бути герб або номінал. При цьому кажуть, що випав герб або номінал. Однозначно передбачити результат підкидання монети неможливо, хоча закони механіки принципово дають нам таку можливість. Однак для цього потрібно з такою точністю задати початкове положення монети, прикладене зусилля, конвективні потоки повітря, мікроскопічні нерівності підлоги і т.п., шо на практиці це здійснити неможливо. Виникає питання: чи можемо ми у такому експерименті знайти певну закономірність? Виявляється, що незважаючи на непередбачуваність результату кожного окремого підкидання, у серіях підкидань проявляється певна об'єктивна закономірність. Вона полягає в тому, що частоти появи герба і номіналу у різних досить довгих серіях приблизно постійні і при зростанні довжини серії все більше наближаються до числа 0.5.

Бюффон у XVIII столітті виконав 4040 підкидань монети; герб випав 2048 раз, і частоти появи герба та номіналу у цій серії виявилися рівними 0.508 та 0.492 відповідно. Значно пізніше Пірсон виконав 24000 підкидань монети; герб випав 12012 раз, так що частоти появи герба та номіналу склали відповідно 0.5005 та 0.4995.

Якби ми провели серію експериментів з підкиданням грального кубика, то переконалися б, що частоти появи кожного із шести чисел, які нумерують грані, з ростом довжини серії наближаються до 1/6.

Виявляється, що для певного кола експериментів описане явище має загальний характер. *Частота появи* певного результату у серії експериментів наближається до деякого числа p при зростанні кількості експериментів у серії.

Факт *стабільності частот* є об'єктивним законом природи для таких явищ. Він служить фундаментом для побудови теорії ймовірностей як математичної дисципліни і одночасно є тією ланкою, яка єднає теорію з практикою.

Означення 1. Назвемо *ймовірнісним* (стохастичним) експериментом такий експеримент, який має певну чітко окреслену множину наслідків (результатів), причому в результаті кожної конкретної реалізації експерименту наступає один і тільки один із них. Ці наслідки називають *елементарними подіями* або *елементарними наслідками* ймовірнісного експерименту. Кожна елементарна подія *повинна* відзначатися *стабільністю частоти* у вище описаному розумінні.

Наприклад, для експерименту, який полягає у підкиданні монети маємо два елементарні наслідки: появу герба і появу номіналу (тут, як і при вивченні будь-якого природного явища, абстрагуємося від деяких неістотних можливостей: нехтуємо можливістю падіння монети на ребро, зникнення у щілині піллоги і т.п.); підкидання грального кубика налічує шість елементарних подій: появу одного із шести перших натуральних чисел.

Підкреслимо ще одну істотну рису ймовірнісних експериментів: вони *повинні бути здійсненними*, принаймні принципово, при *незмінній сукупності основних визначальних умов необмежене* число раз. У протилежному випадку не має змісту поняття стабільності частоти елементарної події і пропадає з'єднуюча ланка між теорією і практикою. Тому теорія ймовірностей не займається вивченням унікальних, одиничних явищ, які проте й належать до категорії таких, що можуть мати різні наслідки. Наприклад, таке висловлювання як: «У 2021 році буде встановлено надійний і постійний контакт з позапланетною цивілізацією», має імовірнісний характер у тому розумінні, що у свій час буде підтверджене або спростоване, проте в силу своєї неповторюваності не може розглядатися в рамках теорії ймовірностей. Розглянемо ще одну ситуацію. Нехай завод виготовляє деякі вироби, перевірка надійності яких полягає в тому, що новий виріб повинен безвідмовно працювати протягом певного часу. Експеримент над кожним конкретним виробом є унікальним і не може бути повторений за незмінних умовах: виріб або перестає бути новим, або виходить з ладу. Однак, як показує практика, при налагодженому масовому виробництві частка надійних (або ненадійних) виробів має властивість стабільності. Тому тут під серією експериментів природно розуміти не повторні випробування одного і того ж об'єкта, а випробування певної кількості однотипних об'єктів. При такій постановці проблема підлягає математичному описові методами теорії ймовірностей.

§3. Ймовірнісний простір як модель стохастичного експерименту

Перейдемо до побудови математичної моделі ймовірнісного експерименту. Як завжди у таких випадках, в основу теорії закладають певні положення, що є узагальненнями досвіду, а дальший розвиток відбувається на, основі методу дедукції. Історично шлях до побудови моделі ймовірнісного експерименту був далеко не простим. Однак підсумки багатовікових зусиль найбільших умів, будучи сформульовані і підкріплені прикладами, легко сприймаються і є досить переконливими. Ми підемо уже второваним шляхом.

§3.1. Простір елементарних подій

Задля простоти спочатку будемо розглядати тільки так звані *дискретні ймовірнісні (стохастичні) експерименти*, які мають скінченну або не більш, ніж зліченну, кількість елементарних наслідків ω. Нагадаємо, що в результаті реалізації експерименту наступає один і тільки один із наслідків. Наслідками імовірнісних експериментів можуть бути числа, кольори, різні предмети і т.п. Тому зрозуміло, що моделлю сукупності елементарних наслідків будь-якого дискретного експерименту може служити дискретна множина

$$\Omega = \{\omega_1, \omega_2, \ldots\},$$

елементи якої «кодують» елементарні наслідки.

Означення 2. Множину Ω називають простором *елементарних подій.*

Проілюструємо цю схему прикладами.

Приклад 1.1. Підкидання монети.

$$\Omega = \{\text{Г}, \text{Н}\}$$

де букви Г і Н позначають відповідно елементарні події, що полягають у появі герба або номіналу.

Приклад 1.2. Підкидання двох монет (або двократне підкидання монети).

$$\Omega = \{\text{ГГ}, \text{ГН}, \text{НГ}, \text{НН}\}.$$

Тут, наприклад, ГН означає появу герба на першій монеті і номіналу на другій (або появу герба при першому підкиданні і появу номіналу при другому підкиданні).

Приклад 1.3. Підкидання грального кубика.

$$\Xi = \{1, 2, 3, 4, 5, 6\}.$$

Число k ($k = \overline{1, 6}$) означає, що випала грань за номером k, тобто випало k очок.

Приклад 1.4. Підкидання тетраедра, грані якого пофарбовані відповідно у білий, жовтий, червоний та синій кольори.

$$\Omega = \{\text{Б,Ж,Ч,С}\}$$

Тут, наприклад, означає, що тетраедр упав на площину гранню, пофарбованою у білий колір.

Приклад 1.5. Підкидання двох гральних кубиків (або двократне підкидання кубика).

$$\Omega = \{(1; 1), (1; 2), (1; 3), \dots, (6; 6)\}.$$

Простір елементарних подій містить 36 елементів вигляду $(i; j)$, де $i, j = \overline{1, 6}$. Елемент $(i; j)$ означає, що випало i очок на першому кубику (при першому підкиданні) та j очок на другому кубику (при другому підкиданні).

Приклад 1.6. Підкидання монети до першої появи герба.

$$\Omega = \{\text{Г}, \text{ГН}, \text{ННГ}, \text{НННГ}, \dots\}.$$

Символ $\underbrace{\text{Н} \dots \text{Н}}_{n-1}\text{Г}$ означає, що герб вперше випав при n-му підкиданні монети. Маємо безліч елементарних подій, оскільки номінал може випадати поспіль довільну кількість раз. (Інтуїція підказує нам, що поява поспіль великої кількості номіналів малоймовірне, однак це не має відношення до принципової можливості такого випадіння).

§3.2. Алгебра подій

У контексті кожного ймовірнісного експерименту можна цікавитися подіями, які не обов'язково співпадають з елементарними. Наприклад, для підкидання кубика можна розглядати подію A: випала парна кількість очок. Коли наступатиме подія A? Очевидно, кожен раз, коли результатом буде поява двох, чотирьох або шести очок. Цей приклад мотивує таке загальне означення.

Означення 3. *Випадковою подією* A називається довільна підмножина A множини Ω. При цьому вважається, шо подія A наступає тоді і тільки тоді, коли результатом ймовірнісного експерименту є одна із елементарних подій, яка належить підмножині A.

Таке означення випадкової події дозволяє трактувати елементарні події, як нерозкладні наслідки випадкового експерименту.

Розглянемо декілька прикладів випадкових подій.

Приклад 1.7. Підкидання грального кубика.

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Подія A — випадає парна кількість очок:

$$A = \{2, 4, 6\}.$$

Приклад 1.8. Двократне підкидання монети.

$$\Omega = \{\text{ГГ}, \text{ГН}, \text{НГ}, \text{НН}\}.$$

Подія A — принаймні один раз випадає герб:

$$A = \{\text{ГГ}, \text{ГН}, \text{НГ}\}.$$

Приклад 1.9. Двократне підкидання кубика.

$$\Omega = \{(i; j);\ i, j = \overline{1, 6}\}.$$

Подія A — сума очок, що випали, дорівнює 4:

$$A = \{(1; 3), (2; 2), (3; 1)\}.$$

Саму множину Ω називають *достовірною подією*: вона наступає у кожній реалізації випадкового експерименту.

Для зручності вводиться *неможлива подія* $\varnothing$, яка відповідає порожній множині — множині, що не містить жодного елемента.

Таким чином, сукупність усіх подій, які можна розглядати для заданого дискретного ймовірнісного експерименту, співпадає з класом усіх підмножин простору елементарних подій Ω. Позначимо цю сукупність через $\mathfrak{A}$. За означенням до $\mathfrak{A}$ належить як достовірна, так і неможлива подія.

Приклад 1.10. Перерахуємо усі можливі події для випадкового експерименту підкидання монети:

$$\Omega = \{\text{Г}, \text{Н}\}, \quad \mathfrak{A} = \{\varnothing, \{\text{Г}\}, \{\text{Н}\}, \Omega\}$$

Оскільки самі події є множинами, над ними можна виконувати теоретико-множинні операції об'єднання, перерізу і доповнення.

Означення 4. Події A і B називають *еквівалентними*, якщо вони можуть відбутися (або не відбутися) тільки разом, тобто $A \subset B$ і $B \subset A$. Позначають $A = B$.

Означення 5. *Сумою подій* A і B називається подія $A + B$ ($A \cup B$), що складається з тих і тільки тих елементарних подій, які належать події A або події B:

$$A + B = \{\omega \in \Omega \colon (\omega \in A) \vee (\omega \in B)\}.$$

Подія $A + B$ наступає тоді і тільки тоді, коли наступає хоча б одна із подій A, B (або A, або B, або A і B).

Означення 6. *Добутком подій* A, B називається подія $A \cdot B$ ($A \cap B$), що складається з тих і тільки тих елементарних подій, які одночасно належать і події A і події B:

$$A \cdot B = \{\omega \in \Omega \colon (\omega \in A) \wedge (\omega \in B)\}.$$

Подія $A \cdot B$ наступає тоді і тільки тоді, коли одночасно наступають події A, B.

Події A, B називаються *несумісними*, якщо їх добуток є неможливою подією: $A \cdot B = \varnothing$.

Означення 7. *Протилежною* до події A називається подія $\overline{A}$, що складається з тих і тільки тих елементарних подій, які не належать події A:

$$\overline{A} = \{\omega \in \Omega \colon \omega \notin A\}.$$

Зауважимо, що

$$A + \overline{A} = \Omega.$$

Означення 8. *Різницею* подій A і B називають подію $A - B$ (A/B), що складається з тих і тільки тих елементарних подій, які належать події A і не належать події B:

$$A/B = A - B = \{\omega \in \Omega \colon (\omega \in A) \wedge (\omega \notin B)\}.$$

Операції об'єднання та перетину подій володіють властивостями комутативності та асоціативності (множина подій є алгеброю Буля):

$$A \cup B = B \cup A, \quad A + B = B + A,$$

$$(A \cup B) \cup C = A \cup (B \cup C) = (A \cup C) \cup B = A \cup B \cup C,$$

$$A \cap B = B \cap A, \quad A \cdot B = B \cdot A,$$

$$(A \cap B) \cap C = A \cap (B \cap C) = (A \cap C) \cap B = A \cap B \cap C,$$

$$(A + B) + C = A + (B + C) = (A + C) + B = A + B + C,$$

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

Об'єднання та перетин подій є дистрибутивними операціями:

$$(A \cup B)C = (AC) \cup (BC),$$

$$(AB) \cup C = (A \cup C)(B \cup C),$$

$$(A \cap B)C = (AC) \cap (BC),$$

$$(AB) \cap C = (A \cap C)(B \cap C).$$

Крім того, слід відмітити, що для всіх $A, B \in \Omega$, виконуються властивості заперечення (закони Моргана)

$$\overline{A \cup B} = \overline{A \cdot B}, \quad \overline{A \cdot B} = \overline{A} \cup \overline{B},$$

які можна узагальнити наступним чином

$$\overline{\bigcup_{i=1}^{k} A_i} = \bigcap_{i=1}^{k} \overline{A_i}, \quad \overline{\bigcap_{i=1}^{k} A_i} = \bigcup_{i=1}^{k} \overline{A_i}.$$

Із властивостей операцій перетину та об'єднання подій можна отримати формулу розкладу випадкової події A на дві несумісні події:

$$A = AB \cup A\overline{B}.$$

Дійсно,

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

І наостанок введемо поняття повної групи подій.

Означення 9. Сукупність подій $A_1, A_2, \ldots, A_n$ називають *повною групою подій*, якщо хоча би одна з них обов'язково відбудеться в результаті експерименту, тобто

$$\bigcup_{k=1}^{n} A_k = \Omega.$$

Означення 10. Сукупність подій $\{A_k\}$, $k = \overline{1, n}$ називають *сприятливими* для події A, якщо виконуються наступні умови:

а) $\bigcup\limits_{k=1}^{n} A_k \subset A$;

б) $A_i \cdot A_j = \varnothing$ для $\forall i, j = \overline{1, n}$ $(i \neq j)$.

В якості систем подій слід розглядати такі класи множин, які називають алгебрами. Нагадаємо, що систему множин $\mathfrak{A}$ називають *алгеброю*, якщо:

1) $\Omega \in \mathfrak{A}$,

2) якщо $A \in \mathfrak{A}$, $B \in \mathfrak{A}$, то $A \cup B$, AB, $A \backslash B \in \mathfrak{A}$, $\overline{A} \in \mathfrak{A}$.

У теорії множин показано, що у випадку, коли простір Ω. має скінченну кількість елементів (елементарних подій), рівну N, то загальне число множин (подій) $N(\mathfrak{A})$, яке складає систему множин $\mathfrak{A}$, є $N(\mathfrak{A}) = 2^n$.

§4. Ймовірність випадкової події

Зробимо наступний крок, а саме припишемо кожній елементарній події $\omega_i \in \Omega$, $i = 1, \ldots, N$ деяку числову функцію $P(\omega_i)$, що задовольняє наступним вимогам:

$$\begin{aligned} &1) \quad P(\omega_i) \geqslant 0, \\ &2) \quad \sum_{i=1}^{N} P(\omega_i) = 1, \\ &3) \quad P(\omega_1 + \omega_2) = P(\omega_1) + P(\omega_2) \end{aligned} \tag{1.1}$$

і означимо *ймовірність* появи випадкової події A

$$A = \bigcup_{i=1}^{k} \omega_i, \quad \omega_i \in A$$

за формулою

$$\mathbb{P}(A) = \sum_{\{i\colon \omega_i \in A\}} P(\omega_i). \tag{1.2}$$

Числову функцію $\mathbb{P}(A)$ випадкової події A, означену формулою (1.2), називають *ймовірністю появи події A*.

Будемо говорити, що трійка $(\Omega, \mathfrak{A}, \mathbb{P})$ задає *ймовірнісну модель (ймовірнісний простір) ймовірнісного (стохастичного) експерименту зі скінченним простором елементарних наслідків Ω та алгеброю подій $\mathfrak{A}$*.

Які властивості ймовірностей слідують з формули (1.2)?

I. $P(\Omega) = 1$ — слідує з означення (1.1) (властивість 2).

II. $P(A) \leqslant 1$ — довести самостійно.

III. Нехай A і B — випадкові події. Тоді справедлива *теорема додавання*

$$P(A + B) = P(A) + P(B) - P(A \cdot B). \tag{1.3}$$

Доведення. Дійсно, подію $A + B$ можна подати через суму несумісних подій $A \cdot \overline{B}$, $\overline{A} \cdot B$, $A \cdot B$. Тобто

$$A + B = A \cdot \overline{B} + \overline{A} \cdot B + A \cdot B.$$

Оскільки $A \cdot \overline{B} \cap B \cdot \overline{A} = \varnothing$, $\overline{A} \cdot B \cap \overline{B} \cdot A = \varnothing$, $A \cdot B \cap A \cdot \overline{B} = \varnothing$, то за властивістю 3

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

З іншого боку $A = A \cdot B + \overline{B} \cdot A$ і $B = B \cdot A + B \cdot \overline{A}$. Тому

$$P(A) = P(A \cdot B) + P(A \cdot \overline{B}), \quad P(B) = P(A \cdot B) + P(B \cdot \overline{A}).$$

Це дає можливість знайти

$$P(A \cdot \overline{B}) = P(A) - P(A \cdot B), \quad P(B \cdot \overline{A}) = P(B) - P(A \cdot B).$$

Отже, маємо:

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

■

Як наслідок з властивостей I та II отримуємо властивість IV.

IV.

$$P(\varnothing) = 0. \tag{1.4}$$

Доведення. Дійсно, $\Omega = \Omega + \varnothing$, причому $\Omega \cap \varnothing = \varnothing$. Тому з (1.3) маємо, що

$$1 = P(\Omega) = P(\Omega) + P(\varnothing).$$

звідси слідує, що $P(\varnothing) = 0$. ■

V. Нехай $\overline{A}$ — подія, протилежна до події A. Тоді

$$P(\overline{A}) = 1 - P(A). \tag{1.5}$$

Доведення. Дійсно, так як $\overline{A} + A = \Omega$, то $P(A + \overline{A}) = P(\overline{A}) + P(A) = 1$. Тому $P(\overline{A}) = 1 - P(A)$. Більше того, $P(\overline{A}) \leqslant 1$. ■

VI. Якщо $B \subset A$, тобто подія A є наслідком події B, то

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

Доведення. Дійсно, подію A можна подати так: $A = (A-B)+B$ (довести самостійно). Оскільки $(A-B)\cap B = \varnothing$, то $P(A) = P(A-B)+P(B)$. Звідси $P(A-B) = P(A)-P(B)$. Так, як $P(A-B) \geqslant 0$, то очевидно, що $P(B) \leqslant P(A)$. ■

VII. Якщо $C = \bigcup\limits_{n=1}^{N} A_n$, а A_n — довільні випадкові події і $\forall n$, $(n \in \mathbb{R})\colon\ A_n \in \mathbb{R}$, то

$$\begin{aligned} P(C) = P\Big(\bigcup_{n=1}^{N} A_n\Big) = \sum_{n=1}^{N} P(A_n) - \sum_{1\leqslant n_1\leqslant n_2\leqslant N}^{N} P(A_{n_1}A_{n_2}) + \\ + \sum_{1\leqslant n_1\leqslant n_2\leqslant n_2\leqslant N}^{N} P(A_{n_1}A_{n_2}A_{n_3}) + \ldots + \\ + (-1)^{N-1} P(A_1 \ldots A_n) \end{aligned} \tag{1.6}$$

Доведення. Доведення проведемо методом математичної індукції. Оскільки для випадку $N = 2$ формула (1.6) справджується (див. властивість III), то припустимо, що (1.6) справджується для випадку $N = m > 2$. Доведемо її для випадку $N = m + 1$. Маємо:

$$\mathbb{P}\left(\bigcup_{n=1}^{m+1} A_n\right) = \mathbb{P}\left(\Big(\bigcup_{n=1}^{m} A_n\Big) \bigcup A_{m+1}\right) =$$

$$= \mathbb{P}\left(\bigcup_{n=1}^{m} A_n\right) + \mathbb{P}(A_{m+1}) - \mathbb{P}\left(\bigcup_{n=1}^{m} \big(A_n \cap A_{m+1}\big)\right) =$$

$$= \Bigg\{ \sum_{n=1}^{m} \mathbb{P}(A_n) - \sum_{1 \leqslant n_1 < n_2 \leqslant m}^{m} \mathbb{P}(A_{n_1} \cap A_{n_2}) +$$

$$+ \ldots + (-1)^{m-1}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_m) \Bigg\} +$$

$$+ \mathbb{P}(A_{m+1}) - \Bigg\{ \sum_{n=1}^{m+1} \mathbb{P}(A_n \cap A_{m+1}) -$$

$$- \sum_{1 \leqslant n_1 < n_2 \leqslant m}^{m} \mathbb{P}(A_{n_1} \cap A_{n_2} \cap A_{m+1}) +$$

$$+ \ldots + (-1)^{m-1}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{m+1}) \Bigg\} =$$

$$= \sum_{n=1}^{m+1} \mathbb{P}(A_n) - \sum_{1 \leqslant n_1 < n_2 \leqslant m}^{m} \mathbb{P}(A_{n_1} \cap A_{n_2}) +$$

$$+ \sum_{1 \leqslant n_1 < n_2 < n_3 \leqslant m+1} \mathbb{P}(A_{n_1} \cap A_{n_2} \cap A_{n_3}) +$$

$$+ (-1)^m \mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{m+1}).$$

При виведенні ми скористалися справедливістю формули (1.6) для розрахунку $\mathbb{P}\big(\bigcup\limits_{n=1}^{m} (A_n \cap A_{m+1})\big)$ та властивостями операцій об'єднання та перетину множин.

Оскільки формула (1.6) справедлива для $N = m + 1$, то вона справедлива для довільних N. ■

Розглянемо деякі приклади побудови ймовірнісного простору.

Приклад 1.11. Для двократного підкидання симетричної монети (або підкидання двох симетричних монет) природно запропонувати наступну модель (див. приклад 2)

Ω	ГГ	ГН	НГ	НН
P	1/4	1/4	1/4	1/4

(1.7)

Експеримент підтверджує правильність цієї моделі.

Виходячи з властивості II, знайдемо ймовірності наступних подій; A – подія, яка полягає в тому, що принаймні один раз випаде герб, тоді

$$P(A) = P(\{\text{ГГ}\}) + P(\{\text{ГН}\}) + P(\{\text{НГ}\}) = 3/4.$$

Аналогічно одержимо для випадку, коли хоча би раз випаде номінал (подія B):

$$P(B) = /4.$$

Так, як $A + B = \Omega$, то $P(A + B) = 1$, тому використовуючи теорему додавання (1.3), знаходимо, що $P(A \cdot B) = 1/2$.

З іншого боку $A \cdot B = \{\text{ГН}\} + \{\text{НГ}\}$, тому $(A \cdot B) = 1/4 + 1/4 = = 1/2$. Оскільки подія $A \cdot B$ означає випадіння монет різними сторонами, то останній результат можна трактувати наступним чином: при підкиданні двох симетричних монет частота випадіння монет різними сторонами приблизно рівна 0.5 (тобто відбувається приблизно в половині усіх випробувань).

Приклад 1.12. Побудуємо простір елементарних подій — усіх можливих наслідків експерименту, який полягає у розміщенні двох куль по двох ящиках.

Кожне з розміщень є нерозкладним наслідком експерименту, тобто елементарною подією. Тому простір елементарних подій Ω можна записати так:

$$\Omega = \{(ab|-), (a|b), (b|a), (-|ab)\} ,$$

де подія $(ab|-)$ означає, що в першому ящику є дві кулі a і b, а в другому — жодної. Якщо кулі a і b різняться між собою, то природно припустити, що всі елементарні наслідки рівноймовірні, тобто кожен з них може відбутися з імовірністю 1/4.

Ω	$(ab\vert -)$	$(a\vert b)$	$(b\vert a)$	$(-\vert ab)$
P	1/4	1/4	1/4	1/4

(1.8)

Нехай подія A полягає в тому, що хоча би в одному ящику є дві кулі, тоді

$$A = \{(ab|-), (-|ab)\}$$

і

$$P(A) = 1/4 + 1/4 = 1/2.$$

Розглянемо тепер випадок, коли кулі a і b відрізнити не можна. Простір елементарних наслідків в такому випадку звузиться:

$$\Omega = \{(ab|-), (a|b), (-|ab)\},$$

а ймовірності появи цих елементарних наслідків є такими:

Ω	$(ab\vert -)$	$(a\vert b)$	$(-\vert ab)$
P	1/4	1/2	1/4

(1.9)

Історично такий підхід до побудови імовірнісного простору для розміщень куль не викликав жодних сумнівів і був покладений в основу побудови статистики Максвела.

Проте можна побудувати іншу модель (статистика Бозе-Ейнштейна)

Ω	$(ab\vert -)$	$(a\vert b)$	$(-\vert ab)$
P	1/3	1/3	1/3

(1.10)

підтвердженням якої є квантово-механічні експерименти.

Якщо порівняти цей імовірнісний простір з моделлю (1.8), то видно, що для тотожніх куль у моделі (1.10) ймовірність «збиратися в одному ящику» більша, ніж у моделі (1.9). Цей ефект своєрідного «притягання» дозволяє пояснити, наприклад, таке фізичне явище, як надплинність гелію.

Якщо розглянути модель імовірнісного простору для випадку, коли існує заборона двом тотожнім кулям збиратися в одному ящику (принцип заборони Паулі), то простір Ω формально матиме наступний вигляд:

$$\begin{array}{c|c} \Omega & (a|b) \\ \hline P & 1 \end{array} \tag{1.11}$$

що разюче відрізняється від моделі (1.9). Модель (1.11), яка підтверджується фізичними експериментами, називають статистикою Фермі-Дірака.

З усього сказаного слід зробити висновок: єдиним критерієм правильності вибраної для опису деякого явища імовірнісної моделі — є практика.

§5. Класична модель. Елементи комбінаторики

§5.1. Класична модель

Означення 11. *Класичною моделлю* називається такий ймовірнісний простір $(\Omega, \mathfrak{A}, P)$, у якому простір елементарних подій Ω скінченний:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$$

і всі елементарні події ω_i є рівноймовірними:

$$P(\omega_i) = p_i = \frac{1}{N}, \quad i = \overline{1, N}.$$

Якщо позначити кількість елементів множини A через $N(A)$, то з формули (1.2) отримуємо, що у класичній моделі ймовірність події A задається рівністю

$$P(A) = \frac{N(A)}{N} = \frac{|A|}{|\Omega|} \tag{1.12}$$

Перші кроки теорії ймовірностей були пов'язані з цією простою схемою (тому вона й названа класичною), тобто з підрахунком N та $N(A)$ для різних конкретних ситуацій. Проте минуло багато часу, перш ніж вдалося чітко сформулювати таке поняття як ймовірність події у вигляді (1.12) (Я. Бернуллі, Монмор, Муавр; початок XVIII ст.).

У попередніх параграфах ми вже застосовували класичну модель, розглядаючи експерименти з підкиданням двох монет та заповненням двох урн двома кулями.

Класична модель описує ті експерименти, які мають скінченну кількість елементарних наслідків, причому умови проведення експерименту забезпечують рівноправність усіх елементарних подій.

Класична модель охоплює величезну кількість практичних задач. Простота класичної схеми насправді оманлива: знайти величини N та $N(A)$ у багатьох ситуаціях далеко не просто. Для цього, в основному, застосовуються комбінаторні методи, до розгляду яких ми переходимо,

§5.2. Основний принцип комбінаторики

При обчисленні кількості різних комбінацій часто використовується, так званий, основний принцип комбінаторики.

При r послідовних виборах з n_k $(1 \leqslant k \leqslant r)$ можливими наслідками одержується $n_1 \cdot n_2 \cdot \ldots \cdot n_r$ різних результатів.

Доведення. Нехай спочатку $r = 2$. Закодуємо n_1 наслідків першого вибору символами $a_1, a_2, \ldots, a_{n_1}$, а n_2 наслідків другого вибору символами $b_1, b_2, \ldots, b_{n_2}$ і утворимо матрицю, у якій на перетині i-го рядка та j-го стовпця стоятиме пара (a_i, b_j). У такій матриці кожна пара зустрічається тільки один раз, а загальна кількість таких пар дорівнює кількості елементів матриці, тобто $n_1 \cdot n_2$, що й стверджує основний принцип комбінаторики.

Тепер застосуємо метод повної математичної індукції.

Нехай результат справедливий для $r = m$. Послідовність з $m + 1$ виборів можна розглядати як два послідовні вибори, з яких перший має $n_1 \cdot n_2 \cdot \ldots \cdot n_m$ можливих наслідків, а другий — n_{m+1} наслідків. За доведеним для $r = 2$ одержимо, що для послідовності з $m + 1$ виборів кількість наслідків дорівнює $(n_1 \cdot n_2 \cdot \ldots \cdot n_m) \cdot n_{m+1} = n_1 \cdot n_2 \cdot \ldots \cdot n_{m+1}$, що й завершує доведення. ■

§5.3. Вибірки

Нехай задано множину M з n довільних елементів:

$$M = \{a_1, a_2, \ldots, a_n\}.$$

Вибіркою без повернення об'єму r $(1 \leqslant r \leqslant n)$ з множини M називається впорядкована послідовність довільних r елементів множини $M = \{a_{i_1}, a_{i_2}, \ldots, a_{i_r}\}$. Такі вибірки можна утворювати наступним чином: вибираємо довільний елемент a_{i_1} з множини M, потім — довільний елемент a_{i_2} з множини $M - a_{i_1}$ і т.д. r разів.

Позначимо кількість вибірок без повернення об'єму r з n елементів через A_n^r. Застосувавши основний принцип комбінаторики до описаного процесу виборів, одержимо

$$A_n^r = n(n-1)(n-2)\ldots(n-r+1) = \frac{n!}{(n-r)!}. \tag{1.13}$$

Величини A_n^r називають також *числами розміщень з* n *по* r.

При $n = r$ одержуємо вибірку об'єму n з n, яка називається *перестановкою.* Число таких вибірок

$$A_n^n = n(n-1)(n-2)\ldots 2 \cdot 1 \equiv n!.$$

Ця величина називається *числом перестановок з* n *елементів* і позначається P_n:

$$P_n \equiv n! \tag{1.14}$$

Зрозуміло, що P_n співпадає з кількістю способів, якими можна впорядкувати множину з n елементів.

Усі комбінаторні результати корисно вміти інтерпретувати «мовою урн і куль».

За самою побудовою A_n^r співпадає з кількістю елементарних наслідків випадкового експерименту, який полягає у вийманні r куль з урн, що містять n занумерованих куль.

Можна підійти і з іншого боку. Уявімо собі, що ми маємо r занумерованих куль і n занумерованих урн. Для першої кулі вибираємо деяку урну і опускаємо у неї кулю; для другої кулі вибираємо одну з порожніх урн і т.д. Кількість наслідків такого експерименту, очевидно, дорівнює A_n^r. Іншими словами, A_n^r співпадає з кількістю розміщень r занумерованих куль у n занумерованих урнах, при яких у кожній урні є не більше, ніж одна куля.

Вибірку можна утворити й іншим способом. З множини M виберемо деякий елемент, запам'ятаємо його і знову повернемо у множину і так r разів. Тепер число r може бути і більшим, ніж n.

Вибіркою з поверненням називається послідовність із r елементів, яка одержується при такому виборі. Очевидно, що вибірка з поверненням може містити і однакові елементи. Позначимо *кількість вибірок з поверненням* через B_n^r. Застосовуючи основний принцип комбінаторики, отримуємо

$$B_n^r = \underbrace{n \cdot n \cdot \ldots \cdot n}_{r} = n^r. \tag{1.15}$$

B_n^r співпадає з кількістю способів, якими можна розмістити r занумерованих куль у n занумерованих урнах без обмежень на кількість куль у кожній урні.

§5.4. Біноміальні коефіцієнти

Розглянемо знову множину M і задамося питанням, скільки різних підмножин по r елементів кожна можна вибрати з множини M (при цьому, очевидно, є зміст розглядати тільки такі r, які не перевищують n). Для побудови підмножин можна використати вибірку без повернення. Однак вибірки без повернення упорядковані, тобто дві вибірки, які складаються з тих самих елементів і відрізняються тільки порядком елементів, вважаються різними. У підмножинах порядок елементів неістотний. Тому підмножин, що містять по r елементів кожна, буде у стільки разів менше, ніж відповідних вибірок, скількома способами можна впорядкувати множину з r елементів. Якщо ми позначимо через C_N^r кількість названих підмножин, то

$$C_n^r = \frac{A_n^r}{r!} = \frac{n!}{r!(n-r)!} \tag{1.16}$$

Числа C_n^r називаються *біноміальними коефіцієнтами*, оскільки вони фігурують у записі бінома Ньютона:

$$(p+q)^n = \sum_{r=0}^{n} C_n^r p^r q^{n-r}, \tag{1.17}$$

тут для простоти запису під 0! розуміємо 1, оскільки з формули (1.16) для $r = n$ маємо

$$C_n^n = \frac{A_n^n}{n!} = \frac{n!}{n!} = 1 = \frac{n!}{n!0!} = \frac{1}{0!} = 1.$$

Тому виходить, що $0! = 1$.

C_n^r ще називають *числом сполук (комбінацій) з n по r*.

З формули (1.17) при $p + q = 1$ одержуємо рівність:

$$2^n = \sum_{r=0}^{n} C_n^r, \tag{1.18}$$

яка доводить, що кількість усіх можливих підмножин заданої множини з n елементів дорівнює 2^n (з урахуванням порожньої підмножини, якій у (1.18) відповідає доданок при $r = 0$, і самої заданої множини — доданок при $r = n$).

Без доведення наведемо ще одну корисну рівність для біноміальних коефіцієнтів:

$$\sum_{r=0}^{n} (C_n^r)^2 = C_{2n}^r. \tag{1.19}$$

§5.5. Поліноміальні коефіцієнти

Нехай, як і раніше, множина M складається з n елементів. Розглянемо фіксований набір цілих чисел k_i, $i = \overline{1,r}$, $1 \leqslant r \leqslant n$, таких, що $0 \leqslant k_i \leqslant n$ і $k_1 + k_2 + \ldots + k_r = n$. Нехай A_i, $i = \overline{1,n}$ — підмножини множини M, які попарно не перетинаються: $A_i \cdot A_j = \varnothing$ при $i \neq j$, причому $|A_i| = k_i$, $i = \overline{1,n}$. За таких умов, очевидно, що

$$M = A_1 + A_2 + \ldots + A_r. \tag{1.20}$$

Задамося питанням, скількома способами можна подати множину M у вигляді (1.20). Це питання еквівалентне наступному: скількома способами n занумерованих куль можна розмістити у r занумерованих урнах за умови, що у першій урні буде k_1 куль, у другій k_2 куль і т.д., а числа k_i задовольнятимуть вищезаписані умови. Число таких способів позначається як $C_n(k_1, k_2, \ldots, k_r)$ називається *поліноміальним коефіцієнтом.*

Оскільки для першої урни k_1 куль можна вибрати $C_n^{k_1}$ способами, після чого для другої урни k_2 куль — $C_{n-k_1}^{k_2}$ способами і т.д., то застосувавши основний принцип комбінаторики, отримуємо:

$$C_n(k_1, k_2, \ldots, k_r) = C_n^{k_1} \cdot C_{n-k_1}^{k_2} \cdot \ldots \cdot C_{n-k_1-k_2-\ldots-k_{r-1}}^{k_r}.$$

Після спрощення з використанням формули (1.16) маємо

$$C_n(k_1, k_2, \ldots, k_r) = \frac{n!}{k_1!k_2!\ldots k_r!}. \quad (1.21)$$

При $r = 2$ поліноміальні коефіцієнти співпадають з біноміальними:

$$C_n(k, n-k) = C_n^k.$$

Також має місце тотожність

$$(z_1 + z_2 + \ldots + z_r)^n = \sum_{\substack{k_1,\ldots,k_r \geqslant 0 \\ k_1+\ldots+k_r=n}} C_n(k_1, \ldots, k_r) z_1^{k_1} \ldots z_r^{k_r}, \quad (1.22)$$

яка узагальнює формулу (1.17).

Поклавши у формулі (1.22) $z_l = z_2 = \ldots = z_r = 1$, отримуємо, що

$$\sum_{\substack{k_1,\ldots,k_r \geqslant 0 \\ k_1+\ldots+k_r=n}} C_n(k_1, \ldots, k_r) = r^n. \quad (1.23)$$

Рівність (1.23) дає розбиття загальної кількості розміщень r куль у n урнах на розміщення з фіксованою кількістю куль у кожній з урн.

§5.6. Формула Стірлінга

Часто виявляється корисною формула, яку ми дамо без доведення і яка називається формулою Стірлінга:

$$n! \sim \sqrt{2\pi n} \cdot n^n \cdot \mathrm{e}^{-n}, \tag{1.24}$$

де під значком «$\sim$» розуміється асимптотична рівність:

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n} \cdot n^n \cdot \mathrm{e}^{-n}} = 1.$$

Наближення Стірлінга на диво точне навіть для малих n: воно дає 0.9221 для 1!, 1.919 для 2!, 118.019 для 5!. Відносні похибки цих наближень відповідно дорівнюють 8%, 4% і 2%.

Приклад 1.13. Знайти ймовірність того, що при випадковому опусканні n куль у r урн всі урни будуть заповнені.

Загальна кількість елементарних подій у такому експерименті дорівнює $B_n^n = n^n$. Оскільки кулі опускають навмання, є підстави вважати, що усі елементарні наслідки рівноймовірні, тому застосуємо класичну схему. Нехай подія A полягає в тому, що всі урни заповнені. Тоді

$$N(A) = P_n = n!$$

Звідси

$$P(A) = \frac{n!}{n^n}$$

При великих n можна використати формулу Стірлінга (1.24), що дає

$$P(A) = \sqrt{2\pi n}\mathrm{e}^{-n}.$$

Ця ймовірність дуже швидко спадає з ростом n. (У певному розумінні випадковість «не терпить рівномірності»).

Приклад 1.14. У групі є n студентів. Чому дорівнює ймовірність того, що принаймні у двох із них збігаються дні народження? Яким повинно бути n, щоб згадана ймовірність була більша, ніж 1/2?

Припустимо, що дні народження людей розподіляються рівномірно за всіма днями року, і кожен рік має 365 днів. Таке припущення дасть нам змогу застосовувати класичну модель. Уявімо собі проведення нашого випадкового експерименту наступним чином. Є 365 урн, занумерованих днями року і n куль, на кожній з яких записано прізвище студента. Куля опускається в ту урну, яка відповідає дню народження студента. Усіх елементарних наслідків такого експерименту є 365^n, і згідно з нашими припущеннями всі наслідки рівноймовірні. Нас цікавить імовірність того, що принаймні в одній урні виявиться більше, ніж одна куля.

Позначимо подію, про яку йдеться, через A. Очевидно, що

$$P(A) = 1, \quad \text{якщо} \quad n > 365.$$

Якщо $n \leqslant 365$, то розглянемо протилежну подію $\overline{A}$, яка полягає в тому, що у кожній урні є не більше, ніж одна куля. Оскільки $N(\overline{A}) = A_{365}^n$, то на основі формули (1.5), одержимо

$$P(A) = \begin{cases} 1, & n > 365, \\ 1 - \dfrac{A_{365}^n}{365^n}, & n \leqslant 365. \end{cases}$$

Якщо $n \ll 365$, то вираз для ймовірності можна спростити:

$$\begin{aligned} P(A) &= 1 - \frac{365!}{(365-n)!365^n} = \\ &= 1 - \frac{365(365-1)\ldots(365-n+1)}{365^n} = \\ &= 1 - \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\ldots\left(1 - \frac{n-1}{365}\right) \approx \\ &\approx 1 - 1 + \frac{1+2+\ldots+(n-1)}{365} = \frac{n(n-1)}{730}. \end{aligned}$$

Нехай $P(A) > 1/2$. Використавши наближену формулу, одержимо нерівність для n

$$\frac{n(n-1)}{365} > 1,$$

розв'язком якої є $n \geqslant 20$.

На основі точної формули одержується оцінка $n \geqslant 23$.

Результат дещо несподіваний: серед великої кількості груп, кожна з яких складається з 23 або більше осіб, більш, ніж у половині груп слід очікувати наявність принаймні двох осіб, дні народження яких співпадають.

Приклад 1.15. Знайти ймовірність того, що серед n вибраних навмання цифр:

а) немає цифри 1;

б) немає цифри 2;

в) немає цифри 1 і 2;

г) немає цифри 1 або цифри 2.

Результатом експерименту є вибірка з поверненням об'єму n з десяти цифр $0, 1, 2, \ldots, 9$. Всього таких вибірок нараховується B_{10}^n і всі вони за умовою задачі рівноймовірні, отже, застосовна класична схема. Події, названі в пунктах а, б, в, г позначимо відповідно A, B, C, D.

Очевидно, що $N(A) = B_9^n = 9^n$. Тому $P(A) = (9/10)^n$.

Аналогічно $P(B) = (9/10)^N$.

Застосовуючи ці ж міркування, знайдемо

$$N(C) = B_8^n = 8^n, \quad P(C) = (8/10)^n.$$

Варто зауважити, що $C = A \cdot B$, однак

$$P(C) \neq P(A) \cdot P(B).$$

За означенням суми подій

$$D = A + B$$

і на основі формули (1.3)

$$P(D) = P(A) + P(B) - P(A \cdot B) = P(A) + P(B) - P(C) =$$
$$= 2 \cdot (9/10)^n - (8/10)^n = \frac{2 \cdot 9^n - 8^n}{10^n}.$$

§6. Аксіоматична побудова теорії ймовірностей. Загальна імовірнісна модель

§6.1. Аксіоми теорії ймовірностей

У попередніх параграфах ми детально розглянули дискретні випадкові експерименти і дискретні імовірнісні моделі. Однак легко навести приклади експериментів, для яких мислима більш, ніж зліченна кількість елементарних наслідків. Наприклад, якщо навмання вибирається деяке число з відрізку $[0, 1]$, то множина елементарних наслідків утворює континуум, і наслідки не можуть бути занумеровані.

Тому для недискретних просторів елементарних подій виникає ситуація, коли не кожна підмножина множини Ω може бути названа подією, оскільки не кожній такій підмножині можна приписати ймовірність. Виявляється, що подібні труднощі мають загальний характер.

Осмислення всіх попередніх результатів разом з бажанням побудувати модель, що описувала б якомога ширше коло ймовірнісних експериментів, веде до необхідності аксіоматичної побудови теорії ймовірностей, яку вперше здійснив видатний російський математик А. М. Колмогоров.

Слід зауважити, що аксіоматичний підхід характерний для кожної розвиненої точної науки. При такому підході в основу теорії кладуться аксіоми, які узагальнюють як практичний досвід, так і досвід роботи з найпростішими моделями; далі теорія розвивається на основі дедукції.

Означення 12. *Загальною ймовірнісною моделлю* (загальним ймовірнісним простором) називається трійка $(\Omega, \mathfrak{A}, P)$, компоненти якої визначаються наступною системою аксіом.

Аксіома I. Ω — довільна множина, яка називається *простором елементарних подій* (наслідків); її елементи називаються *елементарними подіями* (наслідками).

Аксіома II. $\mathfrak{A}$ — деякий клас підмножин множини Ω, який називається *алгеброю подій* і задовольняє наступні умови:

1) $\Omega \in \mathfrak{A}$, $\varnothing \in \mathfrak{A}$;

2) з того, що $A \in \mathfrak{A}$, випливає, що $\overline{A} \in \mathfrak{A}$;

3) з того, що $A \in \mathfrak{A}$ і $B \in \mathfrak{A}$ випливає, що $A+B \in \mathfrak{A}$ і $A \cdot B \in \mathfrak{A}$.

Якщо разом з властивостями 1, 2 і 3 додатково виконується умова:

3') з того, що $A_n \in \mathfrak{A}$, $n = 1, 2, \ldots$, випливає, що

$$\bigcup_{n=1}^{\infty} A_n = (A_1 + A_2 + \ldots) \in \mathfrak{A} \quad \text{і} \quad \bigcap_{n=1}^{\infty} A_n = (A_1 \cdot A_2 \cdot \ldots) \in \mathfrak{A},$$

то клас $\mathfrak{A}$ називають *σ-алгеброю* (читається: сигма-алгебра).

Суть полягає у тому, що з умови II. 3 за індукцією можна вивести приналежність до $\mathfrak{A}$ скінченних сум і добутків за умови, що їх компоненти належать до $\mathfrak{A}$. Умова II.3' гарантує це для сум і добутків, побудованих із безлічі компонент. Для розвитку теорії умова II.3' має принципове значення і у цьому ми переконаємося далі.

Елементи алгебри (σ-алгебри) $\mathfrak{A}$ називаються *подіями.* Як і раніше, події Ω і $\varnothing$ називаються відповідно *достовірною* і *неможливою*, подія $\overline{A}$ називається *протилежною* до події A. Якщо $A \cdot B = \varnothing$, то події A і B називаються *несумісними.*

Властивості II. 1–II. 3 (II. 3') називають *аксіомами алгебри (σ-алгебри) подій.*

Аксіома III. 1) Кожній випадковій події A ставиться у відповідність невід'ємне число $P(A)$, яке називається *ймовірністю* цієї *події*; при цьому:

2) $P(\Omega) = 1$;

3) якщо події $A_1, A_2, \ldots, A_n$ $(n \in \mathbb{N})$ попарно несумісні, то

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

(аксіома адитивності ймовірності).

У випадку, коли $\mathfrak{A}$ — σ-алгебра, замість умови 3 можна вимагати виконання сильнішої умови

3') якщо події A_n, $n = 1, 2, \ldots$ попарно несумісні, то

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P\left(A_n\right)$$

(аксіома σ-адитивності ймовірності).

При цьому із аксіом σ-алгебри і властивостей III. 1, III. 2 випливає збіжність ряду $\sum\limits_{n=1}^{\infty} P(A_n)$.

Властивості III. 1–III. 3 називаються аксіомами *ймовірності.*

Три пераховані групи аксіом цілком визначають певний математичний об'єкт, який можна далі вивчати суто математичними методами.

Структура цього об'єкта є достатньо багатою, щоб вмістити все розмаїття реальних фізичних випадкових експериментів. У цю схему вкладаються, звичайно, і вже вивчені нами найпростіші моделі. Однак вибираючи певним чином Ω, $\mathfrak{A}$ і P ми можемо описати і такі експерименти, елементарними наслідками яких є лінії (наприклад, нарисовані самописцем деякого приладу), траєкторії броунівського руху, тощо. Розділ теорії ймовірностей, що вивчає випадкові простори, у яких Ω є певною множиною функцій, називається *теорією випадкових процесів.* Саме аксіоматизація теорії ймовірностей дала можливість для розвитку цієї надзвичайно важливої у прикладному відношенні теорії.

§6.2. Найважливіші наслідки з аксіом

При вивченні дискретного ймовірнісного простору у §4 ми вивели ряд корисних властивостей ймовірностей випадкових подій. Однак цей вивід ґрунтувався на спеціальному заданні ймовірності (формула (1.2)).

Однак при доведенні властивостей ймовірності $\mathbb{P}(A)$, як числової функції означеної на Ω ми ніде не використовували формулу (1.2). Це значить, що всі раніше доведені властивості можна перенести і на аксіоматичний випадок. Ці властивості є такими.

Властивість 1. Якщо $B \subset A$, тобто подія A є наслідком події B (або інакше: подія A наступає кожен раз, коли наступає подія A), то

$$\mathbb{P}(A - B) \equiv \mathbf{P}(A/B) = \mathbb{P}(A) - \mathbb{P}(B). \tag{1.25}$$

Властивість 2. Якщо $B \subset A$, то

$$\mathbb{P}(B) \leqslant \mathbb{P}(A). \tag{1.26}$$

Це слідує з властивості 1.

Властивість 3. Для кожної події $A \subset \Omega$

$$0 \leqslant \mathbb{P}(A) \leqslant 1. \tag{1.27}$$

Доведення. Те, що $\mathbb{P}(A) \geqslant 0$ слідує з аксіоми III. З властивості 2 слідує, що

$$\mathbb{P}(A) \leqslant \mathbb{P}(\Omega) = 1.$$

■

Властивість 4. (Ймовірність протилежної події $\overline{A}$) Для довільної події A

$$\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A) \tag{1.28}$$

Доведення. Для протилежних подій

$$A + \overline{A} = \Omega \quad \text{і} \quad A \cdot \overline{A} = A \cap \overline{A} = \varnothing.$$

Тому за аксіомою III. 2

$$\mathbb{P}(A + \overline{A}) = \mathbb{P}(A) + \mathbb{P}(\overline{A}) = \mathbb{P}(\Omega) = 1,$$

що рівносильно (1.28) ■

Властивість 5. (Теорема додавання) Для довільних подій A і B

$$\mathbb{P}(A+B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \tag{1.29}$$

Доведення аналогічне доведенню, яке викладене раніше.

Варто ще зауважити, що у загальній імовірнісній моделі з того, що $P(A) = 0$ не випливає, що $A = \varnothing$.

Властивість 6. Теореми неперервності.

Теорема 1.1. (неперервність ймовірності знизу). *Нехай $A_1, A_2, \ldots, A_n, \ldots$ — послідовність випадкових подій — така, що $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ і $\bigcap\limits_{n=1}^{\infty} A_n = A$. Тоді*

$$\exists \lim_{n\to\infty} P(A_n) = P(A) = P\left(\bigcap_{n=1}^{\infty} A_n\right). \tag{1.30}$$

Доведення. Подію $A_n, \forall n \in \mathbb{N}$ можна записати наступним чином:

$$A_n = \bigcup_{k=n}^{\infty} (A_k \overline{A}_{k+1}) \bigcup A.$$

Усі доданки в останній рівності попарно несумісні, тому за аксіомою III. 3 отримуємо, що

$$P(A_n) = \sum_{k=n}^{\infty} P(A_k \overline{A}_{k+1}) + P(A).$$

Звідки

$$P(A_n) - P(A) = \sum_{k=n}^{\infty} P(A_k \overline{A}_{k+1}). \tag{1.31}$$

Зокрема, при $n = 1$ з останньої рівності маємо:

$$P(A_1) - P(A) = \sum_{k=1}^{\infty} P(A_k \overline{A}_{k+1}), \tag{1.32}$$

що говорить про збіжність ряду $\sum\limits_{k=1}^{\infty} P(A_k\overline{A}_{k+1})$.

Так, як різниця (1.31) є залишком ряду (1.32), то

$$\exists \lim_{n\to\infty} \big(P(A_n) - P(A)\big) = 0,$$

що означає наступне:

$$\exists \lim_{n\to\infty} P(A_n) = P(A) = P\left(\bigcap_{n=1}^{\infty} A_n\right).$$

■

Теорема 1.2. (неперервність ймовірності зверху). *Нехай $A_1, A_2, \ldots, A_n, \ldots$ — послідовність випадкових подій — така, що $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$ і $\bigcup\limits_{n=1}^{\infty} A_n = A$. Тоді*

$$\exists \lim_{n\to\infty} P(A_n) = P(A) = P\left(\bigcup_{n=1}^{\infty} A_n\right). \tag{1.33}$$

Доведення. Подію A запишемо у вигляді

$$A = \bigcup_{n=1}^{\infty} A_n = A_1 + (A_2 - A_1) + (A_3 - A_2) + \ldots.$$

Оскільки для $\forall i \in \mathbb{N}$ $A_i \cap (A_{i+1} - A_i) = \varnothing$, то за аксіомою III. 3:

$$P(A) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \\ = \lim_{n\to\infty} \{P(A_1) + P(A_2 - A_1) + P(A_3 - A_2) + \ldots\}.$$

Маємо

$$P(A) = \lim_{n\to\infty} \{P(A_1) + P(A_2 - A_1) + P(A_3 - A_2) + \ldots\} = \\ = \lim_{n\to\infty} P(A_n).$$

Отже,

$$\exists \lim_{n\to\infty} P(A_n) = P(A) = P\left(\bigcup_{n=1}^{\infty} A_n\right).$$

■

Варто зауважити наступне. Аксіоми I–III (аксіоми Колмогорова) є фактично означенням міри, введеної на алгебрі $\mathfrak{A}$ простору Ω. Тому можна дати наступне означення ймовірності.

Означення 13. Нехай $\mathfrak{A}$ — σ-алгебра. Відображення $\mathbb{P}\colon \mathfrak{A}_\sigma \to \mathbb{R}$ називають *зліченно-аддитивною ймовірнісною мірою* на σ-алгебрі $\mathfrak{A}_\sigma$, якщо

1) $\mathbb{P}$ — зліченно-аддитивна міра на $\mathfrak{A}_\sigma$ (виконуються аксіоми III);

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

Зліченно-аддитивну ймовірнісну міру $\mathbb{P}$, яка визначена на $\mathfrak{A}_\sigma$ називають *ймовірністю*.

Пару $(\Omega, \mathfrak{A})$ називатимемо *ймовірнісною моделлю* експерименту.

§7. Модель геометричних ймовірностей

Розглянемо дещо загальнішу ситуацію: нехай множина наслідків випадкового експерименту може бути однозначно співставлена з деякою вимірною областю $\Omega \in \mathbb{R}^n$. Як відомо з курсу функціонального аналізу кожній вимірній множині, можна співставити число, яке називають *мірою* і позначають символом "meas" (від французького слова "measure" — міра.

Прикладами вимірних множин є:

в $\mathbb{R}^1$ — область, для якої визначено поняття довжини;

в $\mathbb{R}^2$ — область, для якої визначено поняття площі;

в $\mathbb{R}^3$ — область, для якої визначено поняття об'єму.

Вимірні множини утворюють σ-алгебру $\mathfrak{A}_\sigma$. Припустимо, що усі елементарні наслідки є рівноможливі. Випадковою подією $A \in \mathfrak{A}_\sigma$ назвемо довільну вимірну підмножину множини Ω. Ймовірність появи цієї події покладемо пропорційною до міри підмножини A, тобто

$$\mathbb{P}(A) = c \cdot \operatorname{meas} A, \tag{1.34}$$

де стала $c > 0$. Введена таким чином ймовірність $\mathbb{P}(A)$ задовільняє всі аксіоми Колмогорова (переконайтесь).

Згідно з аксіомами Колмогорова

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

тому отримуємо, що

$$c = \frac{1}{\operatorname{meas} \Omega}.$$

Це дозволяє (1.34) переписати так

$$\mathbb{P}(A) = \frac{\operatorname{meas} A}{\operatorname{meas} \Omega}. \tag{1.35}$$

Модель (1.35) називають *моделлю геометричних ймовірностей.*

Таке означення ймовірності можна наочно проілюструвати таким прикладом: в область Ω навмання кидають точку і цікавляться ймовірністю того, що вона влучить у підмножину A області Ω. Слово «навмання» трактується таким чином: кинута точка з однаковою можливістю може влучити у будь-яку точку області Ω, а ймовірність того, що вона влучить у підмножину A, пропорційна мірі A і не залежить ні від положення підмножини A відносно області Ω, ні від її форми.

Модель геометричних ймовірностей (1.35) нагадує класичну модель, але має і багато істотних відмінностей.

Пам'ятаємо, що у класичній моделі тільки неможлива подія мала нульову ймовірність. У геометричній моделі кожна елементарна подія (точка) має нульову ймовірність, оскільки міра точки і на прямій, і на площині, і у просторі дорівнює нулю. Нульову ймовірність мають і дискретні множини точок, і лінії — на площині, у просторі — лінії і поверхні, і т.д. Строго додатну імовірність мають, образно кажучи, «суцільні» області тієї ж розмірності, що й Ω. Простір елементарних подій Ω містить і невимірні підмножини, для яких поняття ймовірності не визначене. Однак ці парадоксальні з першого погляду факти не стоять на заваді великій практичній значущості геометричної моделі. Розглянемо декілька прикладів.

Приклад 1.16. Коефіцієнти p і q квадратного рівняння

$$x^2 + px + q = 0$$

вибирають навмання з проміжку $[0, 1]$. Чому дорівнює ймовірність того, що корені рівняння будуть дійсними?

Результатом експерименту є впорядкована пара чисел (p, q), кожне з яких належить відрізку $[0, i]$, тобто Ω можна ототожнити з квадратом на площині з декартовими координатами p, q (див. рис. 1.1).

Корені рівняння будуть дійсними, якщо дискримінант невід'ємний:

$$D = p^2 - 4q \geqslant 0,$$

тобто

$$q \leqslant \frac{p^2}{4}.$$

Остання нерівність визначає заштриховану частину квадрата, яка лежить під кривою $q = p^2/4$ і відповідає події A, ймовірність якої шукається.

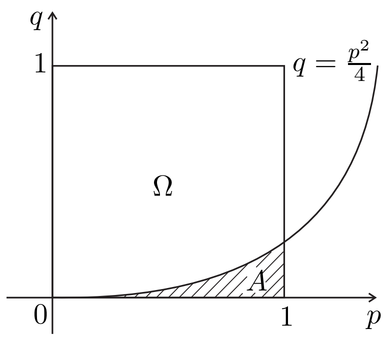

Рис. 1.1.

Оскільки $\operatorname{meas} \Omega = 1$, то

$$P(A) = \operatorname{meas} A = \int_0^1 \frac{p^2}{4}\,\mathrm{d}p = \frac{1}{12}.$$

Приклад 1.17. Стержень розламали навмання на три частини. Яка ймовірність того, що з утворених частин можна скласти трикутник?

Позначимо довжини утворених частин (зліва направо) через x, y, z. Якщо загальну довжину стержня позначити через l, то множиною елементарних наслідків експерименту Ω будуть трійки чисел (x, y, z), що задовольняють умови

$$\begin{cases} x + y + z = l, \\ x \geqslant 0,\ y \geqslant 0,\ z \geqslant 0. \end{cases}$$

Очевидно, що Ω ототожнюється з частиною площини $x + y + z = l$, розташованою у першому октанті (див. рис. 1.2).

Трикутник з відрізків довжини x, y, z можна утворити тоді і тільки тоді, коли довжина однієї сторони не більша за суму двох інших, тобто коли виконуються умови

$$\begin{cases} x \leqslant y + z, \\ y \leqslant x + z, \\ z \leqslant x + y. \end{cases}$$

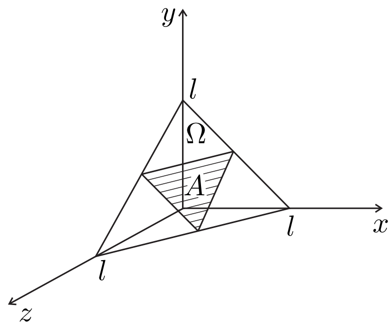

Рис. 1.2.

Додавши до обох сторін першої нерівності x, другої — y і третьої — z та використавши рівняння площини, одержимо умови, що визначають можливість утворення трикутника (подію A):

$$\begin{cases} 0 \leqslant x \leqslant l/2, \\ 0 \leqslant y \leqslant l/2, \\ 0 \leqslant z \leqslant l/2. \end{cases}$$

Легко переконатися, що ці умови задають на Ω трикутник, вершини якого лежать на серединах сторін правильного трикутника Ω. Умови задачі дозволяють застосувати геометричну модель. При цьому

$$P(A) = \frac{\operatorname{meas} A}{\operatorname{meas} \Omega} = \frac{1}{4},$$

оскільки A є одним із чотирьох однакових трикутників, що складають Ω (див. рис. 1.2).

§8. Поняття умовної ймовірності. Ймовірність добутку подій

§8.1. Означення умовної ймовірності

Інтуїтивно зрозуміло, що між двома подіями може існувати зв'язок, взаємний вплив: події можуть мати менш або більш яскраво виражену тенденцію відбуватися (або не відбуватися) одночасно. Крайніми проявами такого взаємозв'язку є несумісність подій та випадок, коли одна із подій є наслідком іншої. Несумісні події ніколи не відбуваються одночасно, а у випадку, коли подія A є наслідком події B, тобто $B \subset A$, подія A відбувається кожен раз, коли відбувається подія B. Для кількісного опису ступеня взаємозалежності подій вводиться поняття *умовної ймовірності.* Розглянемо деякий випадковий експеримент. Виберемо певну подію H назвемо її *гіпотезою* (або *умовою*). Нехай при достатньо великій кількості n повторень експерименту подія H відбулася n_H разів, так що частота її появи у цій серії $\nu_H = n_H/n$.

Тепер для будь-якої іншої події A ми можемо поцікавитися, у скількох випадках із n_H випадків, коли відбувалася подія H, одночасно відбувалася і подія A. Позначимо число таких випадків через n_{AH}.

Відношення n_{AH}, очевидно, *кількісно* характеризує взаємозв'язок подій A і H. Це відношення зручно можна записати у вигляді

$$\frac{n_{AH}}{n_H} = \frac{n_{AH}/n}{n_H/n} = \frac{\nu_{AH}}{\nu_H}$$

де в чисельнику ми маємо частоту появи добутку подій A і H, а в знаменнику — частоту появи події H. В силу стабільності частот, їх можна замінити ймовірностями.

Ці навідні міркування мотивують наступне означення.

Означення 14. Нехай $(\Omega, \mathfrak{A}, P)$ — загальний імовірнісний простір, $H \in \mathfrak{A}$ і $P(H) \neq 0$. *Умовною ймовірністю* $P(A|H)$ появи події A за умови, що наступила подія H, називається число

$$P(A|H) = \frac{P(A \cdot H)}{P(H)} \tag{1.36}$$

Для несумісних подій A і H: $A \cap H = \varnothing$ і $P(A|H) = 0$; якщо подія A є наслідком події H, тобто $H \subset A$: $A \cap H = H$, то $P(A|H) = 1$.

У класичній моделі умовна ймовірність може бути записана і в такому вигляді

$$P(A|H) = \frac{N(A \cdot H)}{N(H)}, \tag{1.37}$$

оскільки

$$P(A \cdot H) = \frac{N(A \cdot H)/N(\Omega)}{N(H)/N(\Omega)} = \frac{N(A \cdot H)}{N(H)}$$

Покажемо, що введена формулою (1.36) *умовна ймовірність* задовольняє аксіоми Колмогорова:

1. $P(A|H) > 0$ — очевидно.
2. $P(\Omega) = 1$. Треба довести, що $P(\Omega|H) = 1$.

Доведення. Нехай $A = \Omega$. Оскільки $H \subset \Omega$, то $\Omega \cap H = H$, тому

$$P(\Omega|H) = \frac{P(\Omega \cdot H)}{P(H)} = \frac{P(H)}{P(H)} = 1.$$

■

3. Доведемо аксіому додавання. Нехай A і B — несумісні події, $A \cap B = \varnothing$. Доведемо, що $P[(A + B)|H] = P(A|H) + P(B|H)$.

Доведення. Якщо A і B несумісні події, то події $A \cdot H$ та $B \cdot H$ — теж несумісні.

$$\begin{aligned} P[(A+B)|H] &= \frac{P[(A+B)\cdot H]}{P(H)} = \frac{P(A \cdot H + B \cdot H)}{P(H)} = \\ &= \left\{\text{за третьою аксіомою Колмогорова}\right\} = \\ &= \frac{P(A \cdot H)}{P(H)} + \frac{P(B \cdot H)}{P(H)} = P(A|H) + P(B|H), \end{aligned}$$

що і треба було довести. ■

Таким чином введена умовна ймовірність

$$P(A|H) = \frac{P(A \cdot H)}{P(H)}$$

задовольняє аксіоми Колмогорова, що означає, що це *ймовірність*.

Доведемо тепер, що

$$P(\overline{A}|H) = 1 - P(A|H)$$

Доведення. Дійсно, так як $\overline{A} \cup A = \Omega$, а $P(\Omega \cdot H) = P(H)$, то маємо

$$1 = \frac{P(H)}{P(H)} = \frac{P(H \cdot \Omega)}{P(H)} = \frac{P\big((A + \overline{A})H\big)}{P(H)} =$$
$$= \frac{P(A \cdot H) + P(\overline{A} \cdot H)}{P(H)} = P(A|H) + P(\overline{A}|H),$$

звідси $P(\overline{A}|H) = 1 - P(A|H)$. ■

Таким чином ймовірнісний простір $\{H, \mathfrak{A}_H, P(A|H)\}$ є простором для умовних ймовірностей. Цей простір будується з урахуванням додаткової інформації про експеримент — подія H відбулася, тому простір елементарних наслідків Ω звузився. Іншими словами, ймовірнісний простір $\{H, \mathfrak{A}_H, P(A|H)\}$ можна отримати з простору $\{\Omega, \mathfrak{A}, P\}$ (як підпростір) шляхом звуження σ-алгебри $\mathfrak{A}$ на $\mathfrak{A}_H$ та звуженням міри $\mathbb{P}$ на H.

Приклад 1.18. Розглянемо сім'ї, що мають двоє дітей. Знайти ймовірність того, що в сім'ї обидві дитини хлопчики за умови, що

a) старша дитина хлопчик;

b) принаймні одна дитина — хлопчик.

Простір елементарних наслідків — $\Omega = \{MM, MF, FM, FF\}$. Будемо вважати, що всі наслідки рівноімовірні. Тоді

$$P(MM) = P(MF) = P(FM) = P(FF) = \frac{1}{4}.$$

Нехай подія H полягає в тому, що старша дитина — хлопчик; подія A — молодша дитина — хлопчик. Для пункту a) нас цікавить поява події $C = (A \cdot H)|H$, де $A \cdot H = A \cap H$ — обидві дитини — хлопчики. Для випадку b) нас цікавить поява події $C = (A \cdot H)|(A + H)$.

Для випадку a):

$$P(C) = P(A \cdot H|H) = \frac{P(A \cdot H)}{P(H)} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} = \frac{1}{2};$$

для випадку b):

$$P(C) = P(A \cdot H|(A + H)) = \frac{P(A \cdot H)}{P(A + H)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

§8.2. Теорема множення

З (1.36) одержується формула для ймовірності добутку двох подій

$$P(A \cdot H) = P(H) \cdot P(A|H) \tag{1.38}$$

При $P(A) \neq 0$ можна розглянути умовну ймовірність

$$P(H|A) = \frac{P(A \cdot H)}{P(H)},$$

звідки для ймовірності добутку двох подій одержимо ще один вираз:

$$P(A \cdot H) = P(A) \cdot P(H|A). \tag{1.39}$$

Зауважимо, що рівність (1.38) має зміст і залишається справедливою і у випадку, коли $P(H) = 0$, оскільки тоді $P(A \cdot H) = 0$ (довести самостійно). Аналогічне зауваження можна зробити і щодо рівності (1.39) при $P(A) = 0$.

З рівностей (1.38), (1.39) одержуємо, так звану, теорему множення, згідно з якою

$$P(A \cdot H) = P(H) \cdot P(A|H) = P(A) \cdot P(H|A), \tag{1.40}$$

тобто ймовірність добутку двох подій дорівнює добутку ймовірності однієї з них на умовну ймовірність другої за умови, що перша відбулася.

§8.3. Незалежність подій

У контексті вищесказаного про поняття умовної ймовірності цілком природним є таке означення.

Означення 15. Нехай $P(H) \neq 0$. Подія A називається незалежною від події H, якщо

$$P(A|H) = P(A). \tag{1.41}$$

Природно було б сподіватися, що якщо подія A незалежна від H, то і подія H незалежна від A. Доведемо, що з означення (1.36) і справді випливає така симетрія.

Доведення. Нехай $P(A) \neq 0$ і подія A незалежна від H. Тоді за теоремою множення (1.40)

$$P(H) \cdot P(A|H) = P(A) \cdot P(H|A),$$

звідки, враховуючи (1.41)

$$P(H|A) = \frac{P(H) \cdot P(A|H)}{P(A)} = \frac{P(H) \cdot P(A)}{P(A)} = P(H),$$

що й треба було довести. ■

Оскільки $P(H|A) = P(H)$, то і $P(A|H) = P(A)$ (згідно з теоремою множення (1.40)). Це значить, що якщо подія A не залежить від H, то і подія H не залежить від A. Далі про такі події будемо говорити, що вони *незалежні*.

Теорема 1.3. *Події A і B незалежні, тоді і тільки тоді, коли*

$$P(A \cdot B) = P(A) \cdot P(B) \tag{1.42}$$

Доведення. Необхідність. Нехай події A і B незалежні. Тоді

$$P(A|B) = P(A)$$

і за теоремою множення (1.40).

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

Достатність. Нехай виконується рівність (1.42). Тоді

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

і подія A незалежна від B, а отже A і B незалежні. ■

Рівність (1.42) називається *теоремою множення для незалежних подій.*

§8.4. Попарна незалежність подій та незалежність в сукупності

Розглянемо групу подій $A_1, A_2, \ldots, A_n$, кожна з яких має ненульову ймовірність.

Група подій $A_1, A_2, \ldots, A_n$ називається *попарно незалежною*, якщо незалежні будь-які дві події A_i, A_j, $i \neq j$, $i, j = \overline{1, n}$.

Група подій $A_1, A_2, \ldots, A_n$ називається *незалежною в сукупності*, якщо для будь-яких

$$1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant n,$$

$$P(A_{i_1} \cdot A_{i_2} \cdot \ldots \cdot A_{i_r}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_r}). \qquad (1.43)$$

З самого означення очевидно, що незалежна в сукупності група подій є попарно незалежною (досить у (1.43) покласти $r = 2$). Однак протилежне твердження неправильне.

Приклад 1.19. Нехай проводиться експеримент з підкиданням тетраедра, грані якого пофарбовані наступним чином: одна біла, друга зелена, третя синя, а на четверту нанесено смугами всі три кольори. Нехай події Б, З, С означають відповідно, що грань, на яку падає тетраедр, містить білий, зелений, синій колір.

Доведемо, що події Б, З, С попарно незалежні. Очевидно,

$$P(\text{Б}) = P(\text{З}) = P(\text{С}) = \frac{2}{4} = \frac{1}{2}, \quad P(\text{БЗ}) = P(\text{БС}) = P(\text{ЗС}) = \frac{1}{4}.$$

Отже, $P(\text{БЗ}) = P(\text{Б}) \cdot P(\text{З})$, і подібно для всіх інших пар, що й означає на основі теореми 1.3 попарну незалежність подій Б, З, С.

Однак

$$P(\text{БЗС}) = \frac{1}{4} \neq P(\text{Б}) \cdot P(\text{З}) \cdot P(\text{С}) = \frac{1}{8},$$

а тому події Б, З, С не є незалежними в сукупності.

Приклад 1.20. Підкидають два гральні кубики. Знайти ймовірність того, що сума очок дорівнює 8 (подія A), за умови, що ця сума парна (подія H).

Застосуємо класичну схему і обчислимо $P(A|H)$ за формулою (1.37).

Знайдемо спочатку $N(H)$. Парна сума очок може приймати значення: 2, 4, 6, …, 12. Легко виписати усі елементарні події, що приводять до таких сум. Всього їх є 18, тому $N(H) = 18$.

Оскільки в нашому випадку $A \cdot H = A$, то $N(A \cdot H) = N(A)$, як легко обчислити, $N(A) = 5$.

Тому

$$P(A|H) = \frac{5}{18}.$$

§9. Формула повної ймовірності та формули Байєса

§9.1. Формула повної ймовірності

Формули, які ми виведемо у цьому параграфі, базуються на понятті умовної ймовірності і у багатьох випадках полегшують знаходження ймовірностей подій.

Теорема 1.4. (формула повної ймовірності). *Нехай події $H_1, H_2, \ldots, H_n$ попарно несумісні і кожна з них має відмінну від нуля ймовірність. Нехай подія A задовільняє умову*

$$A \subset (H_1 + H_2 + \ldots + H_n).$$

Тоді

$$P(A) = \sum_{k=1}^{n} P(H_k) P(A|H_k). \tag{1.44}$$

Доведення. Умови теореми дозволяють записати подію A так:

$$A = A \cdot H_1 + A \cdot H_2 + \ldots + A \cdot H_n,$$

причому несумісність подій H_k зумовлює несумісність подій $A \cdot H_k$:

$$(A \cdot H_k) \cap (A \cdot H_m) = \varnothing, \quad \text{якщо} \quad k \neq m.$$

Тому за третьою аксіомою Колмогорова:

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

Застосувавши до кожного з доданків у правій частині останньої рівності теорему множення (1.40), одержимо, що

$$P(A) = \sum_{k=1}^{n} P(H_k) P(A|H_k).$$

■

➢ ***Зауваження 1.1.*** *Події H_k часто називають гіпотезами. Якщо група гіпотез $H_1, H_2, \ldots, H_n$ крім попарної несумісності задовольняє ще й умову $H_1 + \ldots + H_n = \Omega$, то вона називається повною групою подій (гіпотез). Очевидно, що для повної групи подій*

$$\sum_{k=1}^{n} P(H_k) = 1.$$

Якщо $H_1, H_2, \ldots, H_n$ — повна група гіпотез, то рівність (1.44) справедлива для будь-якої події A.

Формула (1.44) широко використовується як у теоретичних викладках, так і при знаходженні ймовірностей подій у конкретних ситуаціях. При цьому важливо звертати увагу на перевірку виконання умов теореми.

Приклад 1.21. В урну, що містить n куль ($n \geqslant 1$), поклали одну білу кулю. Яка ймовірність того, що навмання вийнята з урни куля буде білою, якщо усі припущення щодо початкового складу урни по кольору куль (може бути білим або чорним) — рівноможливі?

Введемо повну групу гіпотез. Нехай гіпотеза H_k полягає в тому, що серед n куль початкового складу урни є точно k білих куль ($0 \leqslant k \leqslant n$). Всього $n+1$ гіпотеза. Сукупність H_k, $k = \overline{0, n}$, утворює повну групу гіпотез.

З рівноймовірності усіх гіпотез випливає, що

$$P(H_k) = \frac{1}{n+1}, \quad k = \overline{0, n}.$$

Нехай подія A полягає в тому, що результатом експерименту є біла куля, тоді

$$P(A|H_k) = \frac{k+1}{n+1},$$

оскільки, припустивши справедливість гіпотези H_k, ми вийматимемо кулю з урни, що містить $n+1$ куль, серед яких $k+1$ білих.

На основі формули повної ймовірності

$$P(A) = \sum_{k=0}^{n} \frac{k+1}{(n+1)^2} = \frac{1}{(n+1)^2} \sum_{k=0}^{n} (k+1) =$$
$$= \frac{1}{(n+1)^2} \frac{n+2}{2}(n+1) = \frac{n+2}{2(n+1)}.$$

§9.2. Формули Байєса

Часто виникає така ситуація: подія A може бути спричинена однією з несумісних подій $H_1, H_2, \ldots, H_n$ (гіпотез). Відомо, що в результаті експерименту подія A наступила. Чому дорівнює ймовірність того, що подія A спричинена гіпотезою H_k $(1 \leqslant k \leqslant n)$?

Точну відповідь на це питання дає наступна теорема.

Теорема 1.5. (формули Байєса). *Нехай справджуються умови теореми 1.3 і* $P(A) \neq 0$. *Тоді*

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{\sum\limits_{l=1}^{n} P(H_l)P(A|H_l)}, \quad k = \overline{1,n}. \tag{1.45}$$

Доведення. За означенням умовної ймовірності (1.36) і теоремою множення (1.40)

$$P(H_k|A) = \frac{P(H_k \cdot A)}{P(A)} = \frac{P(H_k)P(A|H_k)}{P(A)}.$$

Тепер, використавши для $P(A)$ формулу повної ймовірності (1.44), отримуємо формули (1.45). ■

Формулам Байєса (1.45) можна дати наступне тлумачення. Нехай подія A може відбутися в різних умовах, які характеризуються n припущеннями (гіпотезами) $H_1, H_2, \ldots, H_n$. Ймовірності $P(H_k)$, $k = \overline{1,n}$ вважаються відомими до появи події A. Їх називають *апріорними ймовірностями гіпотез* (від латинського "a priori" — до випробування). Умовні ймовірності появи події A при реалізації різних гіпотез $P(A|H_k)$ — також відомі. Здійснюється експеримент, в результаті якого подія A може відбутися або не відбутися. Якщо ж вона відбулася, це дає можливість переоцінити ймовірності реалізації кожної з гіпотез, знайшовши умовні ймовірності $P(H_k|A)$ за формулами Байєса. Ці ймовірності називають *апостеріорними ймовірностями гіпотез* (від латинського "a posteriori" — після випробування).

Приклад 1.22. Прилад складається з двох вузлів: 1 і 2. Ймовірності безвідмовної роботи вузлів 1 і 2 протягом певного часу відповідно дорівнюють $p_1 = 0.8$ і $p_2 = 0.9$. Вузли виходять з ладу незалежно один від одного, і вихід з ладу хоча би одного вузла спричиняє несправність усього приладу. Виявилося, що прилад вийшов з ладу. Знайти ймовірність того, що несправність приладу спричинено виходом з ладу:

a) тільки першого вузла;

b) тільки другого вузла;

c) обох вузлів.

Про роботу приладу протягом заданого часу можна висловити наступні гіпотези:

H_0 — прилад залишається справним;

H_1 — виходить з ладу тільки вузол 1;

H_2 — виходить з ладу тільки вузол 2;

H_3 — виходять з ладу обидва вузли.

Ці гіпотези утворюють повну групу подій.

Нехай подія A полягає в тому, що прилад вийшов з ладу. Тоді $A \cap H_0 = \varnothing$, і можна обмежитися розглядом тільки гіпотез H_1, H_2, H_3. (Іншими словами, має місце не тільки очевидне для повної групи подій включення $A \subset H_0 + H_1 + H_2 + H_3$, а й $A \subset H_1 + H_2 + H_3$).

Ймовірності гіпотез, в силу незалежності вузлів, дорівнюють

$$
\begin{aligned}
P(H_1) &= (1 - p_1)p_2 = 0.18,\\
P(H_2) &= p_1(1 - p_2) = 0.08,\\
P(H_3) &= (1 - p_1)(1 - p_2) = 0.02.
\end{aligned}
$$

Умовні ймовірності за умовою задачі

$$P(A|H_1) = P(A|H_2) = P(A|H_3) = 1.$$

Необхідно знайти a) $P(H_1|A)$ — ?

За формулами Байєса (1.45):

$$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(A)},$$

$$P(A) = 0.18 + 0.08 + 0.02 = 0.28,$$

$$P(H_1|A) = \frac{0.18 \cdot 1}{0.28} = \frac{18}{28} = \frac{9}{14} \approx 0.64.$$

b) $P(H_2|A) = \dfrac{0.08}{0.28} = \dfrac{8}{28} = \dfrac{2}{7} \approx 0.29.$

c) $P(H_3|A) = \dfrac{0.02}{0.28} = \dfrac{1}{14} \approx 0.07.$

Отже, умовні ймовірності гіпотез H_1, H_2, H_3 за умови, що прилад вийшов з ладу, тобто подія A відбулася, $P(H_k|A)$ більші, ніж відповідні апріорні ймовірності $P(H_k)$. Тобто відбулася переоцінка гіпотез з урахуванням отриманого результату експерименту.

Зауважимо, що $P(H_1|A) + P(H_2|A) + P(H_3|A) = 1.$

ЛЕКЦІЯ 2

Послідовні незалежні випробування

§1. Схема Бернуллі. Біноміальна модель

Нехай здійснюються послідовні випробування, в результаті кожного з яких може відбутися або не відбутися подія A. Настання події A надалі будемо називати «успіхом». Ймовірність «успіху» у кожному випробуванні $P(A)$ будемо вважати незалежною від випробування і рівною p, $0 < p < 1$. Такі повторні випробування називають *послідовними незалежними випробуваннями*, або *випробуванням Бернуллі (схемою Бернуллі)*. Прикладом випробування Бернуллі можуть бути послідовні підкидання монети («успіх» — випадання герба); послідовне кидання кубика («успіх» — випадання певної грані), тощо.

У кожному з цих прикладів ймовірність появи «успіху» однакова у будь-якому випробуванні і не залежить від кількості «успіхів» у попередніх випробуваннях ($p = 1/2$ чи $p - 1/6$ відповідно).

Нехай k — число «успіхів» у серії з n послідовних незалежних випробувань, у кожному з яких ймовірність «успіху» дорівнює p. Необхідно знайти ймовірність настання випадкової події, яка полягає в тому, що у серії з n послідовних незалежних випробувань число успіхів дорівнюватиме k. Позначимо цю ймовірність $P_n(k)$.

Покладемо $q = 1 - p$ — це ймовірність «невдачі», тобто ймовірність того, що подія A не відбудеться при одному випробуванні

$P(\overline{A})$. Для знаходження $P_n(k)$ поступимо таким чином.

Нехай — випадкова подія, яка полягає в тому, що в серії з n послідовних незалежних випробувань подія A відбудеться k разів і не відбудеться $n-k$ разів. Наприклад, у перших k випробуваннях подія A відбулася, а в наступних $n-k$ — не відбулася. Тоді

$$B_k = \underbrace{A \cdot A \cdot \ldots \cdot A}_{k} \cdot \underbrace{\overline{A} \cdot \overline{A} \cdot \ldots \cdot \overline{A}}_{n-k}. \qquad (2.1)$$

Так, як випробування незалежні, то за теоремою множення для незалежних подій (1.40):

$$P(B_k) = \underbrace{P(A) \cdot \ldots \cdot P(A)}_{k} \cdot \underbrace{P(\overline{A}) \cdot \ldots \cdot P(\overline{A})}_{n-k} = p^k q^{n-k}. \qquad (2.2)$$

Число таких подій B_k дорівнює числу сполук з n елементів по k елементів, тобто $C_n^k = \frac{n!}{k!(n-k)!}$ $(0! = 1)$.

Так як події B_k попарно несумісні, то за третьою аксіомою Колмогорова шукана ймовірність дорівнює сумі ймовірностей усіх можливих подій B_k. Оскільки $P(B_k)$ однакові, то $P_n(k)$ дорівнює ймовірності однієї події B_k, помноженої на їх число, тобто:

$$P_n(k) = C_n^k p^k q^{n-k}. \qquad (2.3)$$

Формулу (2.3) називають *біноміальною формулою (формулою Бернуллі)*. Розглянемо функцію

$$\varphi_n(u) = (q + p\,u)^n \qquad (2.4)$$

Легко бачити, що коефіцієнт біля u^k є $P_n(k) = C_n^k p^k q^{n-k}$. Таким чином, ймовірність появи події k разів у серії з n незалежних випробувань можна отримати з виразу (2.4), знайшовши коефіцієнт біля u^k. Функцію (2.4) називають *твірною функцією* для ймовірностей $P_n(k)$.

Побудуємо ймовірнісний простір для схеми Бернуллі. Знайдемо

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

Основна аксіома Колмогорова виконується.

Якщо ввести простір елементарних наслідків

$$\Omega = \{\omega\colon \omega = (\lambda_1, \ldots, \lambda_n),\ \lambda_i = 0, 1\},$$

де $\lambda_i = 1$ — «успіх», а $\lambda_i = 0$ — «невдача», то трійка

$$\{\Omega, \mathfrak{A}, P_n(k)\}$$

утворює ймовірнісний простір згідно з аксіомами Колмогорова. Набір чисел $P_n(k)$ називають *біноміальним розподілом*.

Знайдемо ймовірність того, що успіх наступив не більше l разів. Таку ймовірність будемо позначати $P_n(k \leqslant l)$. Нехай подія B полягає в тому, що у серії з n випробувань «успіх» наступив не більше l разів. Тоді $B = A_0 + A_1 + \ldots + A_l$, де A_i — подія, яка полягає в тому, що успіх у серії з n випробувань наступив i разів, $i = \overline{1, l}$. Так як $\bigcap\limits_{i \neq j} A_i A_j = \varnothing$, то

$$P(B) = P(A_0) + P(A_1) + \ldots + P(A_l) = \sum_{i=0}^{l} C_n^i p^i q^{n-i}.$$

Тому шукана ймовірність

$$P_n(k \leqslant l) = \sum_{i=0}^{l} C_n^i p^i q^{n-i}. \tag{2.5}$$

Знайдемо тепер $P_n(k \geqslant l)$ — ймовірність того, що в серії з n випробувань успіх наступив не менше, ніж l разів.

Нехай подія C полягає в тому, що у серії з n випробувань «успіх» наступив більше, ніж l разів. Тоді $C = \overline{B}$, і ймовірність $P(C)$ можна шукати так:

$$P(C) = P_n(k \geqslant l) = 1 - P_n(k \leqslant l) =$$
$$= \sum_{i=0}^{n} C_n^i p^i q^{n-i} - \sum_{i=0}^{l} C_n^i p^i q^{n-i} = \sum_{i=l+1}^{n} C_n^i p^i q^{n-i},$$

тобто шукана ймовірність є

$$P_n(k \geqslant l) = \sum_{i=l+1}^{n} C_n^i p^i q^{n-i} \tag{2.6}$$

З формул (2.5) та (2.6) можна знайти ймовірність того, що у серії з n незалежних випробувань «успіх» відбудеться не менше за k_1 і не більше, ніж k_2 разів.

$$P_n(k_1 \leqslant k \geqslant k_2) = \sum_{i=k_1}^{k_2} C_n^i p^i q^{n-i} = \sum_{i=k_1}^{k_2} P_n(i). \tag{2.7}$$

Приклад 2.1. З умов випуску лотереї відомо, що виграє 1/20 усіх випущених білетів.

а) Скільки білетів треба купити, щоб ймовірність виграшу була не менше 0.99.

б) Яка ймовірність того, що з 200 білетів виграє не менше, ніж 5.

Придбання лотерейних білетів можна розглядати як незалежні випробування, в кожному з яких ймовірність виграшу («успіху») є $p = 1/20 = 0.05$. Тоді ймовірність того, що серед n куплених білетів виграє k, обчислюється за формулою Бернуллі (2.3)

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

а) Очевидно, що ймовірність виграшу — це ймовірність того, що серед n куплених білетів є принаймні один виграшний, тобто $P_n(k \geqslant 1) > 0.99$. Знайдемо n, для яких виконується нерівність

$$P_n(k \geqslant 1) = 1 - P_n(0) = 1 - (1-p)^n = 1 - (0.95)^n \geqslant 0.99$$

$$(0.95)^n \leqslant 0.01$$

$$n \geqslant -\frac{2}{\lg(0.95)} \approx 89.8.$$

Отже, для того, щоби виграти з імовірністю не меншою, ніж 0.99 треба купити принаймні 90 білетів.

б) Тут $n = 200$. Оскільки має виграти не менше 5 квитків, то шукана ймовірність буде

$$P_{200}(k \geqslant 5) = 1 - P_{200}(k < 5) = 1 - \sum_{i=0}^{4} C_{200}^{i} (0.05)^{i} (0.95)^{200-i} \approx 0.974.$$

§2. Найбільш імовірне число «успіхів» у схемі Бернуллі

Дослідимо питання про те, як змінюються ймовірності $P_n(k)$ з ростом k від 0 до n. Для цього підрахуємо відношення

$$\begin{aligned}\frac{P_n(k)}{P_n(k-1)} &= \frac{\frac{n!}{k!(n-k)!} p^k q^{n-k}}{\frac{n!}{(k-1)!(n-k+1)!} p^{k-1} q^{n-k+1}} = \\ &= \frac{(n-k+1)p}{k\,q} = \frac{(n-k+1)p + k\,q - k\,q}{k\,q} = \\ &= 1 + \frac{(n+1)p - k}{k\,q}, \quad k = \overline{1, n}.\end{aligned}$$

Тепер є очевидним, що

$$\begin{aligned} P_n(k) &> P_n(k-1), \quad \text{якщо} \quad k < (n+1)p, \\ P_n(k) &< P_n(k-1), \quad \text{якщо} \quad k > (n+1)p, \\ P_n(k) &= P_n(k-1), \quad \text{якщо} \quad k = (n+1)p, \end{aligned} \tag{2.8}$$

Існує тільки одне ціле число k_0, яке задовольняє умову

$$(n+1)p - 1 < k_0 \leqslant (n+1)p, \tag{2.9}$$

це число

$$k_0 = [(n+1)p]$$

де $[x]$ означає цілу частину x.

Аналізуючи співвідношення (2.8), бачимо, що ймовірності $P_n(k)$ зростають при зміні k від 0 до k_0, після чого спадають при зміні k від k_0 до n. Тому ймовірність $P_n(k)$ має максимальне значення при $k = k_0$; якщо, до того ж, $(n+1)p$ — ціле число, то найбільшу і однакову ймовірність мають два значення: k_0 і $k_0 - 1$.

Підсумуємо: якщо у схемі Бернуллі $(n+1)p$ не є цілим числом, то найбільш імовірне число «успіхів» k_0 знаходиться з нерівності (2.9), тобто k_0 є цілою частиною числа $(n+1)p$:

$$k_0 = (n+1)p;$$

якщо $(n+1)p$ є цілим числом, то найбільш (і однаково) ймовірними є два значення

$$k_0 = (n+1)p \quad \text{і} \quad k_0 - 1.$$

§3. Асимптотична формула Пуассона для схеми Бернуллі

Встановимо одну важливу граничну властивість величин $P_n(k)$

$$\exists \lim_{\substack{n\to\infty \\ p\to 0 \\ np=\lambda=\mathrm{const}}} P_n(k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \quad k = 0, 1, 2, \ldots. \tag{2.10}$$

Дійсно, прості перетворення дають

$$\begin{aligned} P_n(k) &= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \\ &= \frac{(np)^k}{k!} \frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}(1-p)^{n-k} = \\ &= \frac{(np)^k}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{k-1}{n}\right)\left(1 - \frac{np}{n}\right)^{n-k}. \end{aligned}$$

Якщо тепер перейти до границі, спрямувавши n до безмежності, а p до нуля так, щоб добуток np залишався рівним деякій сталій $\lambda > 0$, то усі вирази в дужках, вигляду $(1 - l/n)$, $l = \overline{1, k-1}$, прямуватимуть до 1, а вираз

$$\left(1 - \frac{np}{n}\right)^{n-k} = \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \underset{k\to\infty}{\longrightarrow} \mathrm{e}^{-\lambda}$$

(оскільки перший множник прямує до $\mathrm{e}^{-\lambda}$, а другий — до 1). Таким чином, одержуємо рівність (2.10).

З рівності (2.10) випливає, що у випадку, коли

$$p \ll 1, \quad k \ll n, \quad n\,p = \mathrm{const}$$

$$P_n(k) \approx \frac{(np)^k}{k!}\mathrm{e}^{-np}. \tag{2.11}$$

Формула (2.11) — це наближення Пуассона для моделі Бернуллі.

При великих n і $k \ll n$ формула (2.11) дає досить добрі результати і часто використовується на практиці, оскільки дозволяє уникнути обчислення факторіалів великих чисел.

Пізніше для моделі повторних випробувань ми одержимо й інші важливі граничні теореми.

Формула (2.10) навіює думку розглянути дискретний ймовірнісний простір, у якому

$$\Omega = \{0, 1, 2, \ldots\},$$

$$P(\{k\}) \equiv p_k = \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots, \quad \lambda > 0. \tag{2.12}$$

Коректність такого задання моделі підтверджується виконанням умови нормування для ймовірностей елементарних подій:

$$\sum_{k=0}^{\infty} p_k = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1.$$

Описана модель називається *моделлю Пуассона.*

Цікавим є питання, за яких умов випадкові експерименти описуються моделлю Пуассона?

Розглянемо конкретний приклад. Нехай експеримент полягає у реєстрації продуктів радіоактивного розпаду (наприклад, альфа-частинок) протягом певного проміжку часу T. Розділимо проміжок T на n однакових частин Δ_k $(k = \overline{1, n})$ довжиною $\Delta t = T/n$ кожна. Вважаючи n достатньо великим числом, припустимо, що справджуються такі умови:

1) впродовж кожного часового проміжку Δ_k реєструється не більше, ніж одна частинка, причому ймовірність реєстрації частинки на кожному проміжку Δ_k однакова і пропорційна його довжині: $p = \alpha\Delta t = \alpha T/n$; відповідно ймовірність відсутності частинки протягом Δ_k для всіх k дорівнює $q = 1 - \alpha T/n$. (Точніше кажучи, ми не можемо цілком виключити можливість реєстрації протягом Δ_k двох і більше частинок, однак вважаємо, що ймовірність такої події є нескінченно мала вищого порядку порівняно з $p = \alpha\Delta t$ при $\Delta t \to 0$);

2) ймовірність появи (чи відсутності) частинки впродовж Δ_k не залежить від результатів попередніх спостережень. За таких умов ми маємо, фактично, n незалежних випробувань (кожне на своєму часовому проміжку Δ_k) з двома можливими результатами: реєстрацією частинки — «успіхом» — з ймовірністю p і відсутністю частинки — «невдачею» — з ймовірністю q. Тому ймовірність зареєструвати k частинок протягом часу T задається формулою Бернуллі

$$P_n(k) = C_n^k p^k q^{n-k} = C_n^k \left(\alpha\frac{T}{n}\right)^k \left(1 - \alpha\frac{T}{n}\right)^{n-k}.$$

Припущення про те, що впродовж проміжку Δ_k лічильник реєструє не більше, ніж одну частинку, тим точніше, чим більше n (тобто, чим менше Δt). Тому в останній формулі доцільно перейти до границі при $n \to \infty$. При цьому $p = \alpha T/n \to 0$, $np = n\alpha T/n = \alpha T \equiv \lambda = \text{const}$, і ми приходимо до результату (2.10):

$$\lim_{n\to\infty} P_n(k) = p_k = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \quad \lambda = \alpha T.$$

Наведені вище міркування легко узагальнюються і на інші ситуації. Нехай, наприклад, підраховується кількість подій, які мають місце у деякій області на площині або у просторі, і ця область має міру (площу або об'єм), рівну T. (Підраховується, скажімо, кількість метеоритів, що випадають протягом певного часу на 1 кв. км місячної поверхні; кількість куріпок на 1 га угідь;

кількість родзинок у кексі; кількість пилинок у 1 л повітря). Припускаючи, що для достатньо малої області ймовірність появи відповідної події у цій області пропорційна мірі області, а ймовірність появи двох і більше подій є нескінченно малою вищого порядку малості порівняно з ймовірністю появи однієї події і що події в підобластях, які не перетинаються, незалежні між собою, знову прийдемо до моделі Пуассона (2.12) з параметром λ, пропорційним мірі області T: $\lambda = \alpha T$.

Параметр λ характеризує інтенсивність процесу, який розглядається: ймовірність того, що подія не наступить ні разу, $p_0 = \mathrm{e}^{-\lambda}$, є тим меншою, чим більше значення λ.

Другою апроксимаційною формулою є формула Лапласа, яку ми подамо без доведення.

§4. Локальна теорема Лапласа-Муавра

Нехай у кожному з n незалежних випробувань ймовірність появи події A є однаковою і дорівнює p, і нехай $P_n(k)$ — ймовірність того, що при цих n випробуваннях подія A відбудеться k разів. Тоді справедлива *локальна теорема Лапласа-Муавра.*

Теорема 2.1.(локальна теорема Лапласа-Муавра). *Нехай* $0 < p < 1$. *Тоді*

$$P_n(k) = \frac{1}{\sqrt{n\,p\,q}} \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-x^2/2}\big(1 + \alpha_n(x)\big), \quad x_k = \frac{k - n\,p}{\sqrt{n\,p\,q}},$$

де $\alpha_n(x) \underset{n\to\infty}{\rightrightarrows}$ *для всіх* k, *за яких* $x_k \in [a, b]$, $|a - b| < +\infty$.

З цієї теореми слідує, що

$$P_n(k) \approx \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{n\,p\,q}} \exp\left\{-\frac{1}{2}\left(\frac{k - n\,p}{\sqrt{n\,p\,q}}\right)^2\right\}. \tag{2.13}$$

Цю формулу ще називають *локальною формулою Лапласа-Муавра.* Її отримують з $P_n(k)$, використовуючи формулу Стірлінга для обчислення $n!$.

З локальної теореми Лапласа можна вивести інтегральну теорему Лапласа.

Знайдемо

$$P_n(k_1 \leqslant k \leqslant k_2) = \sum_{i=k_1}^{k_2} C_n^i p^i q^{n-i}.$$

Справедлива наступна теорема.

Теорема 2.2.Інтегральна теорема Лапласа-Муавра. *Нехай μ — число успіхів у серії n незалежних випробувань, в кожному з яких ймовірність успіху дорівнює p. Тоді для довільних a, b $(a < b)$ має місце співвідношення*

$$\lim_{n\to\infty} P\left\{a \leqslant \frac{\mu - n\,p}{\sqrt{n\,p\,q}} \leqslant b\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-x^2/2}\,\mathrm{d}x.$$

Приклад 2.2. Деяке підприємство дає 99.2% стандартних виробів. Яка ймовірність того, що серед 5000 навмання вибраних приладів цього підприємства число нестандартних виробів менше 60? Вибір окремих приладів можна розглядати як послідовні незалежні випробування.

Ймовірність вибрати нестандартний виріб $p = 0.008$; число випробувань $n = 5000$; $q = 1 - \; = 0.992$. Якщо μ — число вибраних нестандартних виробів

$$P(0 \leqslant \mu \leqslant 60) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} \mathrm{e}^{-x^2/2}\,\mathrm{d}x \approx \Phi(x_2) - \Phi(x_1),$$

де

$$x_1 = \frac{k_1 - n\,p}{\sqrt{n\,p\,q}}, \quad x_2 = \frac{k_2 - n\,p}{\sqrt{n\,p\,q}}, \quad k_1 = 0, \quad k_2 = 0,$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \mathrm{e}^{-x^2/2}\,\mathrm{d}x.$$

Тут $\Phi(x)$ — табульована інтегральна функція Муавра-Лапласа.

$$x_1 = -\frac{40}{6.3}, \quad x_2 = \frac{20}{6.3}, \quad \Phi(x_2) - \Phi(x_1) \approx 0.9992.$$

§5. Повторні випробування у випадку змінних умов експерименту

Формули (2.3) та (2.4) легко узагальнити на випадок, коли ймовірність появи випадкової події A приймає різні значення в різних випробуваннях (повторні випробування в змінних умовах). Якщо випробування задовільняють всі вимоги схеми Бернуллі, а ймовірність появи події A в j-му випробуванні дорівнює p_j ($q_j = 1 - p_j$), $j = \overline{1, n}$, то замість формули (2.3) тим самим шляхом отримуємо наступну формулу для розрахунку ймовірності появи події A k разів в серії з n випробувань:

$$P_n(k) = \sum p_{j_1} \dots p_{j_k} q_{j_k+1} \dots q_{j_n} \quad (k = 0, 1, \dots, n). \tag{2.14}$$

В (2.14) сума поширюється на всі можливі розбиття чисел $1, 2, \dots, n$ на дві групи:

– перша містить k чисел $(j_1, \dots, j_k)$;

– друга містить $n - k$ чисел $(j_{k+1}, \dots, j_n)$.

Очевидно, що число таких розбиттів є C_n^k. Ймовірність $P_n(k)$ (формула (2.14)) в цьому випадку є коефіцієнтом при u^k в розкладі за степенями u твірної функції

$$\varphi_n(u) = \prod_{j=1}^{n} (q_j + u\, p_j). \tag{2.15}$$

Легко переконатися (зробити самостійно), що ймовірність $P_n(k)$, означена формулою (2.14) задовільняє аксіоми Колмогорова.

§6. Поліноміальна модель

Міркування, використані при побудові моделі Бернуллі (її ще називають біноміальною моделлю) неважко узагальнити на випадок послідовності незалежних випробувань, кожне з яких має

r елементарних наслідків:

$$\Omega_0 = \{\omega_1, \omega_2, \dots, \omega_r\}, \quad r \geqslant 2,$$

$$P(\{\omega_k\}) = p_k, \quad \sum_{k=1}^{r} p_k = 1.$$

Елементарними наслідками експерименту, який полягає у незалежному повторенні n разів випробування Ω_0, будуть «слова», що складаються з r символів, кожен з яких належить «алфавіту» Ω_0. Таких «слів» буде r^n. Якщо «слово» $\omega_{i_1}\omega_{i_2}\dots\omega_{i_n}$ містить k_1 символів ω_{i_1}, k_2 символів ω_{i_2}, $\dots$, k_r символів ω_{i_r} $(k_1 + k_2 + \dots + k_r = n)$, то

$$P(\{\omega_{i_1}\omega_{i_2}\dots\omega_{i_n}\}) = p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}.$$

Імовірність $P_n(k_1, k_2, \dots, k_r)$ того, що при n незалежних випробуваннях k_1 разів відбудеться подія ω_1, k_2 разів відбудеться подія ω_2, $\dots$, k_r разів відбудеться подія ω_r буде рівна

$$P_n(k_1, k_2, \dots, k_r) = C_n(k_1, k_2, \dots, k_r) p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}, \qquad (2.16)$$

де $C_n(k_1, k_2, \dots, k_r)$ — поліноміальні коефіцієнти (1.21).

Формула (2.16) узагальнює формулу Бернуллі (2.3), одержану для моделі Бернуллі.

Приклад 2.3. Розглянемо ідеальний газ із N молекул, що знаходиться в посудині об'єму V_0. Поділимо уявно посудину на дві частини об'єму V і V', так що $V + V' = V_0$. Розглянемо множину (ансамбль) з великої кількості таких посудин з газом. Нехай p — ймовірність того, що деяка молекула знаходиться в об'ємі V, а q — ймовірність її знаходження в об'ємі V'. Якщо газ знаходиться у рівновазі, то молекули рівномірно розподілені в об'ємі посудини і

$$p = \frac{V}{V_0}, \quad q = \frac{V'}{V_0} \quad (p + q = 1).$$

Чому дорівнює ймовірність $P_N(n)$ того, що в об'ємі V, виявиться n молекул (при цьому $N-n$ молекул будуть знаходитися в об'ємі V)?

Оскільки молекули ідеального газу рухаються незалежно, то застосовна модель Бернуллі: кожна молекула з ймовірністю p може знаходитися в об'ємі V («успіх») і з ймовірністю q в об'ємі V' («невдача»). Тому

$$P_N(n) = C_n^n p^n q^{N-n} = C_N^n \left(\frac{V}{V_0}\right)^n \left(\frac{V'}{V_0}\right)^{N-n}.$$

Інтерпретація цього результату така: частина тих посудин в ансамблі, які містять n молекул в об'ємі V (і $N-n$ молекул в об'ємі V), дорівнює $P_N(n)$.

ЛЕКЦІЯ 3

Випадкові величини та їх опис

§1. Випадкові величини. Типи випадкових величин. Означення випадкової величини

Одним з основних понять теорії ймовірності є поняття *випадкової величини.* Історично поняття *випадкової величини* виникло з розвитком зацікавленості до опису похибок вимірювань. Остання залежить від багатьох причин, впливи яких неможливо врахувати. З погляду розглянутих раніше підходів до опису випадкових подій похибка експерименту є випадковою подією, якій можна приписати те чи інше числове значення. Тобто необхідно говорити про числову функцію, значення якої визначаються результатами деякого випробування (вимірювання). Але результатами випробування ω є елементами (точками) простору елементарних подій Ω. Тому природньо назвати випадковою величиною числову функцію $\xi = \xi(\omega)$, що визначена на просторі елементарних наслідків Ω, Для вияснення додаткових умов, які слід накласти на $\xi(\omega)$, розглянемо деякі приклади.

1. Нехай

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \overline{\in} A \end{cases}$$

— індикатор підмножини $A \in \Omega$. Введемо числову функцію $\xi(\omega) =$

$= I_A(\omega)$. Випадкова величина $\xi(\omega)$ приймає тільки два значення

$$\xi(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \overline{\in} A. \end{cases}$$

2. Нехай випробування (стохастичний експеримент) полягає в двократному підкиданні монети з простором елементарних наслідків

$$\Omega = \{\text{ГГ}, \text{ГН}, \text{НГ}, \text{НН}\},$$

визначимо випадкову величину $\xi(\omega)$ як число "гербів", яке випадає при двократному підкиданні монети. Функцію $\xi(\omega)$ зручно задати за допомогою наступної таблиці

ω	ГГ	НГ	ГН	НН
$\xi(\omega)$	2	1	1	0

3. Експеримент полягає в здійсненні серії з n послідовних незалежних випробувань в схемі Бернуллі. Кожний елементарний наслідок ω такого випробування — це деяка послідовність успіхів та невдач в n випробуваннях. Позначимо через $\xi(\omega)$ число успіхів в серії з n випробувань. $\xi(\omega)$ може набувати будь-якого цілого значення k, $k = \overline{0, n}$.

4. Випробування полягає в послідовному підкиданні монети до першого випадання номіналу. Простір елементарних наслідків Ω у цьому випадку нескінченний (але зліченний!)

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\},$$

де $\omega_1 = \text{Н}$, $\omega_2 = \text{ГН}$, $\omega_3 = \text{ГГН}$, … (Н — випадання номіналу, Г — випадання герба). Нехай випадкова величина $\xi(\omega) = n$ — число підкидань монети до першого випадання номіналу. Можливі значення $\xi(\omega)$ — будь-які натуральні числа, $\xi(\omega) = n$, $n \in \mathbb{N}$.

5. Випробування полягає в тому, що стрілець робить один постріл по мішені. Простір елементарних наслідків Ω — множина всіх точок ω мішені. Для кожної точки $\omega \in \Omega$ позначимо через $\xi = \xi(\omega)$ відстань від точки попадання до центра мішені. Очевидно, що $\xi(\omega)$ набуває будь-якого невід'ємного значення, $\xi(\omega) \in \mathbb{R}^+$.

6. Експеримент полягає у вимірюванні відхилень положення космічного апарату відносно розглядуваної (теоретично) траєкторії польоту. Простір елементарних наслідків Ω для такого експерименту — множина всіх точок $\omega \in \mathbb{R}^3$. Нехай $\xi(\omega) = (a_1(\omega), a_2(\omega), a_3(\omega))$ — відхилення координат космічного апарату від координат точок теоретично розрахованої траєкторії. Зрозуміло, що $\xi(\omega)$ може набувати будь-яких значень.

У наведених прикладах розглянуто ряд випадкових величин — числових функцій, визначених на множині елементарних наслідків Ω. Ці приклади дозволять стверджувати, що відносно можливості набування тих чи інших значень випадкові величини $\xi(\omega)$ можна поділити на три типи:

— *дискретні випадкові величини*;

— *неперервні випадкові величини* ($\xi(\omega)$ приймає *тільки неперервні значення*);

— *випадкові величини мішаного типу* ($\xi(\omega)$ приймає *дискретні* та *неперервні значення*).

Але з теоретико-ймовірнісної точки зору задання тільки самої *числової функції* $\xi(\omega)$ ще недостатньо для характеристики випадкової величини $\xi = \xi(\omega)$, оскільки значення які вона набуває, є випадковими. Треба ще мати змогу відповісти на запитання, пов'язані з ймовірностями появи набутих значень.

В прикладах 1–4 ці ймовірності легко знайти. У першому прикладі $P\{\xi = I_A(\omega)\} = 1$; у другому прикладі теж легко обчислюється, наприклад, $P\{\xi = 2\} = \frac{1}{4}$. В третьому прикладі, очевидно, що $P\{\xi = k\}$ обчислюється за біноміальною формулою. В четвертому прикладі, зрозуміло, що

$$P\{\xi = m\} = \left(\frac{1}{2}\right)^{m-1} \cdot \frac{1}{2} = \frac{1}{2^m}, \quad m \in \mathbb{N}.$$

Однак в прикладах 5–6 значення випадкової величини $\xi(\omega)$ заповнюють цілі проміжки з множини дійсних чисел. Тому треба задати не ймовірності появи окремих значень (окремі значення утворюють множини міри 0!), а ймовірність того, що значення випадкової величини $\xi = \xi(\omega)$ попадуть в певний проміжок. Це означає, що для будь-якої випадкової величини ξ природньо поставити вимогу, щоби $\forall x_1, x_2 \in \mathbb{R}$ було визначено ймовірність того, що $x_1 \leqslant \xi < x_2$. Зокрема, $\forall x \in \mathbb{R}$ необхідно визначити ймовірність того, що $\xi < x$. Цю вимогу і покладають в основу наступного означення поняття випадкової величини.

Означення 1. Нехай $\langle \Omega, \mathfrak{A}, P \rangle$ — ймовірнісний простір. Функцію $g\colon \Omega \to \mathbb{R}$ будемо називати вимірною (відносно пари σ-алгебр $(\mathfrak{A}, \mathbb{B}(\mathbb{R}))$ ($\mathbb{B}(\mathbb{R})$ — алгебра борелівських множин), якщо для кожної борелівської множини її прообраз є подією (елементом з $\mathfrak{A}$), тобто

$$(\forall B \in \mathbb{B}(\mathbb{R}))\colon g^{-1}(B) \in \mathfrak{A}.$$

Тоді справедливе наступне

Означення 2. Довільна вимірна функція $\xi(\omega)$ відносно пари σ-алгебр $(\mathbb{B}(\mathbb{R}), \mathfrak{A})$ називається випадковою величиною.

Тобто ξ — випадкова величина, якщо для довільної борелевої множини $B \in \mathbb{B}(\mathbb{R})$ її прообраз

$$\xi^{-1}(B) \stackrel{\text{def}}{=} \{\omega \in \Omega\colon \xi(\omega) \in B\} \in \mathcal{A}$$

є випадквою подією.

Розглянемо ймовірнісний простір $\langle \Omega, \mathfrak{A}, P \rangle$ і функцію $\xi = \xi(\omega)\colon \Omega \to \mathbb{R}$. Справедливе наступне твердження.

Функція $\xi = \xi(\omega)\colon \Omega \to \mathbb{R}$ є випадковою величиною тоді і тільки тоді, коли

$$(\forall x \in \mathbb{R})\colon \xi^{-1}((-\infty, x)) \stackrel{\text{def}}{=} \{\omega \in \Omega\colon \xi(\omega) < x\} \in \mathfrak{A}.$$

Це означає, що $\forall x \in \mathbb{R}$ виконується умова

$$S = \{\omega \in \Omega\colon \xi(\omega) < x\} \in \mathfrak{A}. \tag{3.1}$$

Множину $\{\omega \in \Omega\colon \xi(\omega) < x\}$ надалі будемо скорочено позначати $\{\xi < x\}$. Аналогічний зміст мають множини $\{x_1 < \xi < x_2\}$, $\{\xi > x\}$.

§2. Функція розподілу випадкової величини. Властивості функції розподілу

Введена виразом (3.1) умова $\{\xi < x\} \in \mathfrak{A}$ означає, що для всіх $x \in \mathbb{R}$ можна визначити ймовірність події S:

$$P\{\xi < x\} = P(S).$$

Означення 3. Введемо функцію

$$F_\xi(x) = P\{\xi < x\}, \quad x \in \mathbb{R}, \tag{3.2}$$

яку будемо надалі називати функцією розподілу випадкової величини ξ.

Простір Ξ значень випадкової величини ξ з σ-алгеброю $\mathcal{B}$ множин в ньому і означеною на цих множинах ймовірністю (3.2) будемо називати ймовірнісним простором $\langle \Xi, \mathbb{B}, \mathcal{F} \rangle$ випадкової величини ξ.

Вияснимо, які властивості має функція розподілу $F_\xi(x)$, означена формулою (3.2)

• область визначення: $x \in \mathbb{R}$;

• область значень $0 \leqslant F_\xi(x) \leqslant 1$ — функція $F_\xi(x)$ додатньо визначена і обмежена;

• оскільки $\forall x_1, x_2 \in \mathbb{R}$, $x_1 < x_2$

$$P(x_1 \leqslant \xi \leqslant x_2) = F(x_2) - F(x_1), \tag{3.3}$$

то $F(x_2) \geqslant F(x_1)$, що означає, що $F_\xi(x)$ є неспадною функцією.

Дійсно, розглянемо подію $\{\xi < x_2\}$, яку можна записати так

$$P\{\xi < x_2\} = P\{\xi < x_1\} + P\{x_1 \leqslant \xi < x_2\},$$

а за означенням (3.2)

$$F_\xi(x_2) = F_\xi(x_1) + P\{x_1 \leqslant \xi < x_2\}.$$

Звідки слідує (3.3). Так, як $P\{x_1 \leqslant \xi < x_2\} \geqslant 0$, то $\forall x_1, x_2 \in \mathbb{R}$

$$F_\xi(x_2) \geqslant F_\xi(x_1).$$

Це означає, що функція розподілу є неспадною функцією.

• Функція розподілу $F_\xi(x)$ неперервна зліва, тобто

$$F_\xi(x_0 - 0) = F(x_0) \tag{3.4}$$

Доведення. Для доведення (3.4) достатньо показати, що

$$(\forall x_0)(\forall (z_n), z_n \uparrow x_0(n \to \infty))\colon \lim_{n\to\infty} F_\xi(z_n) = F_\xi(x_0).$$

Для цього розглянемо послідовність множин

$$S_n = \{\omega\colon \xi(\omega) < z_n\}$$

де $\{z_n\}_1^\infty$ — довільна монотонно зростаюча числова послідовність, що збігається до x_0, тобто $\exists \lim\limits_{n\to+\infty} z_n = x_0$.

Очевидно, що $\forall n \geqslant 1$, $S_n \subset S_{n+1}$ $(S_1 \subset S_2 \subset S_3 \subset \ldots)$, тобто послідовність множин S_n є монотонно зростаючою. Оскільки $\bigcup\limits_{n=1}^{+\infty} S_n \overset{\text{def}}{=} B = \{\omega\colon \xi(\omega) < x_0\}$, то за неперервністю ймовірності (див. формулу (1.32)) маємо

$$\begin{aligned}F_\xi(x_0) &= P\{\omega\colon \xi(\omega) < x_0\} = P(B) = \\ &= P\left\{\bigcup_{n=1}^{+\infty} S_n\right\} = \lim_{n\to+\infty} P(S_n) = \\ &= \lim_{n\to+\infty} F_\xi(z_n).\end{aligned}$$

Тому

$$(\forall x_0)(\forall (z_n), z_n \uparrow x_0(n \to \infty))\colon \lim_{n\to\infty} F_\xi(z_n) = F_\xi(x_0) = F_\xi(x_0 - 0).$$

Властивість доведена. ■

• Доведемо тепер, що

$$P\{\omega \in \Omega\colon \xi(\omega) = x_0\} = F(x_0 + 0) - F(x_0) \tag{3.5}$$

Доведення. Нехай $\{z_n\}_1^\infty$ — довільна спадна числова послідовність така, що $z_n \uparrow x_0 (n \to +\infty)$. Розглянемо монотонно спадну послідовність множин $A_n = \{\omega\colon x_0 \leqslant \xi(\omega) \leqslant z_n\}$, для якої виконуються умови

$$\forall n \geqslant 1, \quad A_{n+1} \subset A_n \quad (A_{n+1} \subset A_n \subset \ldots \subset A_1).$$

Оскільки

$$\bigcap_{n=1}^{\infty} A_n \stackrel{\text{def}}{=} B = \{\omega \in \Omega\colon \xi(\omega) = x_0\},$$

то за неперервністю ймовірності (див. формулу (1.34))

$$\begin{aligned} P(B) = P\{\omega \in \Omega\colon \xi(\omega) = x_0\} = P\left\{\bigcap_{n=1}^{\infty} A_n\right\} = \\ = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} \big(F_\xi(z_n) - F_\xi(x_0)\big) = \\ = \lim_{n\to\infty} F_\xi(z_n) - F_\xi(x_0) = F_\xi(x_0+0) - F_\xi(x_0) \end{aligned}$$

■

З цієї властивості слідує, що якщо x_0 — точка неперервності функції розподілу $F_\xi(x)$, то

$$P\{\omega \in \Omega\colon \xi(\omega) = x_0\} = 0!$$

тобто $P(\xi = x) \neq 0$ *тоді і тільки тоді, коли x є точкою розриву функції розподілу.*

Якщо ж функція розподілу є неперервною, то

$$P\{\xi = x\} = 0!$$

— парадокс ймовірності. Ймовірність неможливої події є нуль. Обернене твердження взагалі кажучи невірне: ймовірність попадання в центр мішені дорівнює нулю, однак ця подія є можливою. Інша проблема полягає в тому, що ймовірність попадання в деяку конкретну точку мішені дорівнює нулю, хоча добрий стрілець завжди попаде в мішень. Це означає, що об'єднання подій, що мають нульову ймовірність може дати подію, що має ймовірність одиницю. Цей парадокс пояснюється різницею між нулем і нескінченно малою. Кожній підмножині деякого інтервалу можна приписати ймовірність так, що ця ймовірність дорівнює нулеві тільки для порожньої підмножини, яка відповідає неможливій події, а для довільної іншої події ймовірність додатня, хоча, можливо і нескінченно мала. Тоді для множини B, якій приписано ймовірність $P(B)$ в традиційному сенсі, отримуємо ймовірність, яка відрізняється від $P(B)$ щонайбільше на нескінченно малу. Отримана нова ймовірність не сигма-адитивна, а лише адитивна. Це значить, що для неперервних випадкових величин $\xi(\omega)$ твердження

$$P\{\omega \in \Omega \colon \xi(\omega) = x\}$$

слід розуміти так:

$$P\{\omega \in \Omega \colon \xi(\omega) = x\} = \lim_{\delta \to 0} P\{\omega \in \Omega \colon x + \delta < \xi(\omega) < x - \delta\}.$$

• Покажемо, що

$$\exists \lim_{x \to -\infty} F(x) = F(-\infty) = 0. \tag{3.6}$$

Доведення. Нехай $\{x_n\}_1^\infty$, $\forall n \in \mathbb{N}$, $x_n \in \mathbb{R}$ спадна послідовність дійсних чисел, $x_1 > x_2 > x_3 > \ldots > x_n > \ldots$ і нехай $\exists \lim\limits_{n \to \infty} x_n = -\infty$. Покладемо $A_n = \{\xi < x_n\}$.

Так як

$$A_1 \supset A_2 \supset A_3 \supset \ldots \supset A_n \supset \ldots$$

і

$$\bigcap_{n=1}^{\infty} A_n = \emptyset,$$

то за теоремою неперервності (див. формул (1.32))

$$\begin{aligned} \lim_{n \to \infty} F_\xi(x_n) = \lim_{n \to \infty} P(\xi < x_n) = \\ = \lim_{n \to \infty} P(A_n) = P(\emptyset) = 0. \end{aligned}$$

Звідси $\exists F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0.$ ■

• Покажемо, що

$$\exists \lim_{x \to +\infty} F(x) = F(+\infty) = 1. \tag{3.7}$$

Доведення. Нехай $\{z_n\}_1^\infty$ — довільна зростаюча числова послідовність дійсних чисел і $\exists \lim\limits_{n \to \infty} z_n = +\infty$.

Розглянемо послідовність множин $A_n = \{\omega\colon \xi(\omega) < x_n\}$ таку, що $\forall n \geqslant 1$, $A_n \subset A_{n+1}$, тобто (A_n) — монотонно зростаюча послідовність множин. Оскільки

$$\bigcup_{n=1}^{+\infty} A_n = \xi^{-1}(\mathbb{R}) = \Omega,$$

то за неперервністю ймовірності (див. формулу (1.32)):

$$\lim_{n \to +\infty} F_\xi(z_n) = \lim_{n \to +\infty} P(A_n) = \\ = \lim_{n \to +\infty} P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = P(\Omega) = 1.$$

Це значить, що

$$\exists \lim_{x \to +\infty} F(x) = F(+\infty) = 1.$$

■

Таким чином, функція розподілу $F_\xi(x) = P\{\omega\colon \xi(\omega) < x\}$ випадкової величини $\xi(\omega)$ є неспадною, неперервною зліва функцією в області визначення $x \in \mathbb{R}$ і задовільняє наступні граничні умови

$$F_\xi(-\infty) = 0, \quad F_\xi(+\infty) = 1.$$

Вказані вище властивості називають *характеристичними властивостями функції розподілу*, оскільки вони виокремлюють з класу всіх дійсних функцій, заданих на всій числовій прямій функції, кожна з яких є функцією розподілу для деякої випадкової величини $\xi(\omega)$.

Справджується наступна теорема:
Нехай $\mathcal{F}(x)$ — функція з $\mathbb{R} \to \mathbb{R}_+$, яка має властивості

1) $\mathcal{F}(x) \in [0, 1]$.
2) $\mathcal{F}(+\infty) = 1$.
3) $\mathcal{F}(-\infty) = 0$.
4) $\mathcal{F}(x)$ — неспадна.
5) $\mathcal{F}(x)$ — неперервна зліва.

Тоді існує ймовірнісний простір $\langle \Omega, \mathfrak{A}, P \rangle$ і існує ξ — випадкова величина на ньому така, що

$$F_\xi(x) \equiv \mathcal{F}(x).$$

На жаль, випадкова величина $\xi(\omega)$ визначається своєю функцією розподілу $F_\xi(x)$ неоднозначно. Різні випадкові величини можуть мати одну і ту ж функцію розподілу. Так, наприклад, нехай ξ набуває значень 0 і 1, причому $P\{\xi = 0\} = P\{\xi = 1\} = \frac{1}{2}$ і нехай $\eta = 1 - \xi$. Очевидно, що $\xi \neq \eta$, однак обидві випадкові величини ξ та η мають одну і ту ж функцію розподілу

$$F_\xi(x) = F_\eta(x) = \begin{cases} 0, & x \leqslant 0, \\ 1/2, & x \in (0, 1], \\ 1, & x > 1. \end{cases}$$

§3. Дискретні випадкові величини. Ряд розподілу. Функція розподілу дискретних випадкових величин

Випадкову величину $\xi(\omega)$ будемо називати *дискретною* випадковою величиною, якщо значення, які вона може набувати, утворюють *скінченну* або *зліченну* множинну. Справедливе наступне твердження:
функція $\xi = \xi(\omega) \colon \Omega \to X$ з не більш, ніж зліченною множиною значень $X = \{x_k \colon 1 \leqslant k \leqslant N\}$ $(N \leqslant +\infty)$ є дискретною випадко-

вою величиною тоді і тільки тоді, коли

$$(\forall k\colon 1 \leqslant k \leqslant N)\colon \{\omega\colon \xi(\omega) = x_k\} \in \mathfrak{A}.$$

Прикладами дискретної випадкової величини є приклади 1–4, які приведені в §1 цього розділу.

Для того, щоб задати випадкову величину досить для кожного з можливих значень x_k випадкової величини x задати ймовірність набування цього значення:

$$P_k = P\{x = x_k\}, \quad k = 1, 2, \ldots .$$

Введемо в розгляд поняття *ряду розподілу* випадкової величини.

Означення 4. Рядом розподілу дискретної випадкової величини x називають таблицю

x	x_1	x_2	$\ldots$
p_k	p_1	p_2	$\ldots$

(3.8)

де в першому рядку розміщено всі можливі значення x_k випадкової величини x, а в другому — ймовірності p_k того, що x набуває ці значення: $p_k = P\{x = x_k\}$. Зрозуміло, що значення $x_1, \ldots, x_k, \ldots$ випадкової величини можуть бути довільними $\forall k \geqslant 1$, $x_k \in \mathbb{R}$: щодо ймовірностей p_k необхідно виконання двох умов:

1) $\forall k$, $p_k \geqslant 0$ (ті значення, для яких $p_k = 0$ можна відкинути).

2) Згідно з аксіомою адитивності

$$P\Big(\bigcup_k \{\xi = x_k\}\Big) = \sum_k p_k = P(\Omega) = 1, \tag{3.9}$$

бо події $\{\xi = x_k\}$ та $\{\xi = x_{k+1}\}$ $(1 \leqslant k \leqslant +\infty)$ попарно несумісні, тобто $\forall k \geqslant 1$, $\{\xi = x_k\} \cap \{\xi = x_{k+1}\} = \varnothing$.

Кожна таблиця (3.8), в якій $\forall k \geqslant 1$, $x_k \in \mathbb{R}$, а $P\{\xi = x_k\} = p_k \geqslant 0$ і $\sum\limits_k p_k = 1$ задає дискретну випадкову величину ξ, для якої вона є рядом розподілу. Справді, досить покласти $\Omega = \{x_1, x_2, \ldots\}$, $P\{\xi = x_k\} = p_k$ і $\xi(x_k) = x_k$.

Розглянемо приклад. Урна містить чотири білі і одну червону кулю. Навмання без повернення виймають з урни по одній кулі, доки не буде вийнята червона. Скласти ряд розподілу випадкової величини x — число вийнятих куль. Які значення приймає x?

$$x_0 = 0, \quad x_1 = 1, \quad x_2 = 2, \quad x_3 = 3, \quad x_4 = 4.$$

Знайдемо ймовірності появи даних значень.

$$\begin{aligned}
p_0 &= P\{x = 0\} = P\{A_0\} = \frac{1}{5},\\
p_1 &= P\{x = 1\} = P\{A_1\} = \frac{4}{5}\cdot\frac{1}{4} = \frac{1}{5},\\
p_2 &= P\{x = 2\} = P\{A_2\} = \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{1}{3} = \frac{1}{5},\\
p_3 &= P\{x = 3\} = P\{A_3\} = \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{2}{3}\cdot\frac{1}{2} = \frac{1}{5},\\
p_4 &= P\{x = 4\} = P\{A_4\} = \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{5}.
\end{aligned}$$

Знайдемо суму

$$\sum_{k=0}^{4} p_k = 5\cdot\frac{1}{5} = 1,$$

значить ряд розподілу є таким

x	0	1	2	3	4
p	1/5	1/5	1/5	1/5	1/5

За заданим рядом розподілу можна побудувати *функцію розподілу* дискретної випадкової величини, скориставшись означенням (3.2):

$$F(x) = P\{X < x\} = \sum_{x_k < x} p_k \tag{3.10}$$

Дійсно, нехай подія A — ξ прийняло значення, яка менше за x, $A = \{\omega\colon \xi(\omega) < x\}$. Якщо подія B_k полягає в тому, що x прийняло значення x_k, $B_k = \{\omega\colon \xi(\omega) < x_k\}$, то $A = \bigcup\limits_{k\colon x_k<x} B_k$, причому $\forall k$, $B_k \cap B_{k+1} = \varnothing$. Тому

$$P(A) = P\{\xi < x\} = F(x) = \sum_{x_k<x} p_k. \tag{3.11}$$

Знайдемо тепер з означення (3.10) $p_k = P\{x = x_k\}$. Так, як $P\{X \leqslant x\} = F(x+0)$, то

$$\begin{aligned} P\{X = x_k\} = P\{X \leqslant x_k\} - P\{X < x_k\} = \\ = F(x_k+0) - F(x_k), \end{aligned} \tag{3.12}$$

тобто значення $p_k = P\{x = x_k\} = \Delta F(x_k)$ дорівнює стрибку функції розподілу в точці $x = x_k$.

Якщо множину значень $\{x_k\}$ випадкової величини x можна впорядкувати в порядку зростання (для скінченної множини значень x_k це можливо зробити), то графік $F(x)$ є східчастою лінією з інтервалами сталості. Нехай для прикладу $\forall k \geqslant 1$, $x_k > 0$, $k = 1 \div N$.

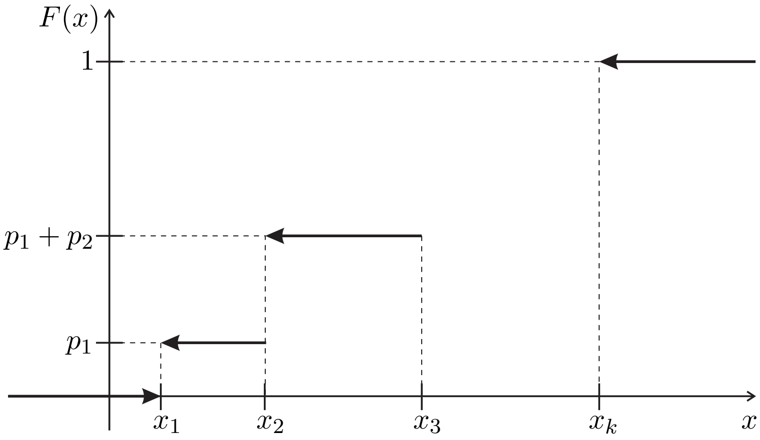

Рис. 3.1. Типовий графік функції розподілу.

Для зліченної множини значень дискретної випадкової величини $\xi(\omega)$ функція розподілу $F_\xi(x) = P\{\xi < x\}$ може і не мати інтервалів сталості. Розглянемо приклад.

Нехай дискретна випадкова величина $\xi(\omega)$ набуває будь-яких значень q з множини раціональних чисел $\mathbb{Q}$ ($q \in \mathbb{Q}$). Пронумеруємо ці значення довільним способом $q_1, q_2, \ldots, q_n, \ldots$. Нехай $P\{\xi = q_k\} = \frac{1}{2^k}$. Оскільки

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = \frac{1}{1-\frac{1}{2}} - 1 = 2 - 1 = 1,$$

то випадкова величина $\xi(\omega)$ має ряд розподілу:

r_1	$\ldots$	r_k	$\ldots$
$\frac{1}{2}$	$\ldots$	$\frac{1}{2^k}$	$\ldots$

Усі раціональні числа є точками розриву 1-го роду для функції розподілу $F(x)$, але графік намалювати не можливо.

§4. Моделі рядів розподілу дискретних випадкових величин

Для задання ряду розподілу дискретної випадкової величини $\xi(\omega)$ слід обов'язково задати ймовірності $p_k = P\{\xi = x_k\}$ $\forall k \geqslant 1$. На практиці дуже часто ці ймовірності моделюють, задаючи спосіб розрахунку ймовірностей p_k.

Найбільш важливими моделями рядів розподілу є:

1. Дискретний рівномірний розподіл. Дискретна випадкова величина $\xi(\omega)$ приймає значення

$$\xi(\omega) = k, \quad k = 1 \div N,$$

ймовірності набування цих значень є

$$p_k = P\{\omega\colon \xi(\omega) = k\} = \frac{1}{N}, \tag{3.13}$$

тут $N = 1, 2, \ldots$ — параметр розподілу.

2. Бернулівський розподіл. Дискретна випадкова величина $\xi(\omega)$ приймає тільки два значення

$$x_1 = 0, \quad x_2 = 1.$$

Ймовірності появи цих значень є

$$\begin{aligned} p_0 &= P\{\omega\colon \xi(\omega) = x_1\} = q \\ p_1 &= P\{\omega\colon \xi(\omega) = x_2\} = p \end{aligned} \tag{3.14}$$

причому

$$p + q = 1.$$

p і q — параметри розподілу, $0 \leqslant p \leqslant 1$, $q = 1 - p$.

3. Біномінальний розподіл. Дискретна випадкова величина $\xi(\omega)$ приймає значення

$$\xi(\omega) = x_k = k, \quad k = 0, 1, \ldots, n.$$

Ймовірності появи цих значень є

$$p_k = P\{\omega\colon \xi(\omega) = k\} = P_k(k) = C_n^k p^k q^{n-k}, \tag{3.15}$$

тут $0 \leqslant p \leqslant 1$, $q = 1 - p$, $n = 0, 1, 2, \ldots$.

4. Пуассонівський розподіл. Дискретна випадкова величина $\xi(\omega)$ набуває цілих додатніх значень

$$\xi(\omega) = k, \quad k = 0, 1, \ldots.$$

Ймовірності появи цих значень

$$p_k = P\{\omega\colon \xi(\omega) = k\} = \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, \tag{3.16}$$

де λ — параметр розподілу. Зауважимо, що умова нормування виконується $\forall \lambda > 0$. Дійсно

$$\begin{aligned} \sum_k p_k = \sum_{k=0}^{\infty} \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!} = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \\ = \mathrm{e}^{-\lambda} \cdot \mathrm{e}^{\lambda} = 1. \end{aligned}$$

5. Геометричний розподіл. Дискретна випадкова величина $\xi(\omega)$ набуває цілих значень

$$\xi(\omega) = k, \quad k \geqslant 1.$$

Ймовірності появи цих значень

$$p_k = P\{\omega\colon \xi(\omega) = k\} = q^{k-1}p, \tag{3.17}$$

тут $0 < p \leqslant 1$, $q = 1 - p$ — параметри розподілу.

§5. Неперервні випадкові величини. Щільність розподілу. Властивості щільності розподілу

Розподіл випадкової величини з незліченною множиною можливих значень неможливо задати ймовірностями окремих значень. Для таких випадкових величин ξ функція розподілу є неперервною, а тому

$$P\{\omega\colon \xi(\omega) = x_0\} = 0, \quad x_0 \in \mathbb{R}.$$

В той же час дуже багато практичних задач пов'язано з такими випадковими величинами (див. приклади 5 та 6 з §1).

Тому для дослідження такого класу випадкових величин використовують той факт, що між ймовірнісними мірами $\mathbb{P}$, заданими на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ та функціями розподілу $F_\xi(x) = P\{\omega\colon \xi(\omega) < x\}$ на числовій прямій $\mathbb{R}$ існує взаємно однозначна відповідність, яка встановлюється рівністю:

$$\mathbb{P}[a, b] = F_\xi(b) - F_\xi(a).$$

Міру $\mathbb{P}$, побудовану на функціях розподілу $F_\xi(x)$ називають *ймовірнісною мірою Лебега-Стільтьеса.*

З відомої в теорії міри теореми Радона-Нікодима випливає, що для такого класу випадкових величин можна ввести поняття *щільності розподілу* $\rho_\xi(x)$. Справедливе наступне означення.

Означення 5. Випадкова величина $\xi(\omega)$ називається *абсолютно неперервною* (або *випадковою величиною зі щільністю*), якщо існує невід'ємна вимірна (за Лебегом) функція $\rho_\xi\colon \mathbb{R} \to \mathbb{R}_+$ така, що

$$(\forall B \in \mathcal{B}(\mathbb{R})) \qquad P\{\omega\colon \xi(\omega) \in B\} = \int\limits_B \rho_\xi(t)\,\mathrm{d}t. \tag{3.18}$$

Функцію $\rho_\xi(x)$ будемо називати *щільністю розподілу* абсолютно неперервної випадкової величини $\xi(\omega)$.

В (3.18) під інтегралом розуміється інтеграл в сенсі Рімана (в загальному випадку це інтеграл Лебега). Поклавши в означенні (3.18) $B = (-\infty, x)$, отримуємо, що

$$F_\xi(x) = \int\limits_{-\infty}^{x} \rho_\xi(t)\,\mathrm{d}t. \tag{3.19}$$

Формула (3.19) дозволяє стверджувати, що майже скрізь (за виключенням множини точок з мірою Лебега нуль)

$$\frac{\mathrm{d}F_\xi(x)}{\mathrm{d}x} = \rho_\xi(x). \tag{3.20}$$

Вивчимо тепер основні властивості щільності розподілу.
1) Безпосередньо з означення (3.19) та властивості неспадності функції розподілу випливає, що $\forall x \in \mathbb{R}$

$$\rho_\xi(x) \geqslant 0.$$

2) Умова нормування:

$$\int\limits_{-\infty}^{\infty} \rho_\xi(t)\,\mathrm{d}t = 1. \tag{3.21}$$

Ця властивість безпосередньо випливає з умови

$$F_\xi(+\infty) = 1.$$

Отож, якщо випадкова величина $\xi(\omega)$ є абсолютно неперервною, то вона може бути описана щільністю розподілу $\rho_\xi(x)$, яка задовільняє властивості 1 та 2. Це значить, що моделювання щільності розподілу $\rho_\xi(x)$ є простішою задачею порівняно з моделюванням функції розподілу $F_\xi(x)$.

Розглянемо декілька прикладів.

Приклад 3.1. Нехай випадкова величина ξ має щільність розподілу

$$\rho_\xi(x) = \frac{A}{1+x^2}, \quad x \in \mathbb{R}.$$

Знайти сталу A.

Згідно з властивістю 1 щільності розподілу $A \geqslant 0$. Оскільки з властивості 2 випливає, що

$$1 = \int\limits_{-\infty}^{\infty} \rho_\xi(t)\,\mathrm{d}t = \int\limits_{-\infty}^{\infty} \frac{A}{1+t^2}\,\mathrm{d}t =$$

$$= A \operatorname{arctg} t \Big|_{-\infty}^{+\infty} = 2A\frac{\pi}{2} = A\pi.$$

Це значить, що $A = \frac{1}{\pi}$. Тому щільність розподілу

$$\rho_\xi(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

Приклад 3.2. На відрізку $[a, b]$ з числової прямої випадково вибирають точку (вважається, що всі положення точки рівноможливі). Нехай значення випадкової величини ξ дорівнює координаті вибраної точки. Знайти функцію розподілу цієї випадкової величини.

З рівноможливості положення точки випливає, що щільність розподілу

$$\rho_\xi(x) = \begin{cases} C, & x \in [a, b], \\ 0, & x \overline{\in} [a, b], \end{cases}$$

(див. означення (3.18)), де стала $C \geqslant 0$. З умови нормування отримуємо, що

$$C = \frac{1}{b-a}.$$

Знайдемо функцію розподілу. З означення (3.19) маємо:
- нехай $x \in]-\infty, a)$, тоді

$$F_\xi(x) = \int\limits_{-\infty}^{x} \rho_\xi(t)\,\mathrm{d}t = \int\limits_{-\infty}^{x} 0\,\mathrm{d}t = 0;$$

- нехай $x \in [a, b]$, тоді

$$F_\xi(x) = \int\limits_{-\infty}^{x} \rho_\xi(t)\,\mathrm{d}t = \int\limits_{-\infty}^{a} \rho_\xi(t)\,\mathrm{d}t + \int\limits_{a}^{x} \rho_\xi(t)\,\mathrm{d}t =$$
$$= \int\limits_{-\infty}^{a} 0\,\mathrm{d}t + \int\limits_{a}^{x} \frac{1}{b-a}\,\mathrm{d}t = \frac{x-a}{b-a};$$

- нехай $x \in (b, +\infty)$, тоді

$$F_\xi(x) = \int\limits_{-\infty}^{x} \rho_\xi(t)\,\mathrm{d}t = \int\limits_{-\infty}^{a} \rho_\xi(t)\,\mathrm{d}t + \int\limits_{a}^{b} \rho_\xi(t)\,\mathrm{d}t + \int\limits_{b}^{+\infty} \rho_\xi(t)\,\mathrm{d}t =$$
$$= \int\limits_{-\infty}^{a} 0\,\mathrm{d}t + \int\limits_{a}^{b} \frac{1}{b-a}\,\mathrm{d}t + \int\limits_{b}^{+\infty} 0\,\mathrm{d}t = \frac{x-a}{b-a} =$$
$$= \frac{b-a}{b-a} = 1.$$

Отож,

$$F_\xi(x) = \begin{cases} 0, & x \in (-\infty, a), \\ \dfrac{x-a}{b-a}, & x \in [a, b], \\ 1, & x \in (b, +\infty). \end{cases}$$

Зазначимо, що означення (3.18) щільності розподілу дозволяє ввести поняття щільності розподілу для дискретних випадкових величин. Нехай $\xi(\omega)$ — дискретна випадкова величина, що приймає значення $x_1, \ldots, x_n$ з відповідними ймовірностями

$$p_i = P\{\omega\colon \xi(\omega) = x_i\}, \quad i = 1 \div N.$$

Згідно з (3.18) для довільного достатньо малого $\varepsilon > 0$

$$p_i = \int\limits_{x_i-\varepsilon}^{x_i+\varepsilon} \rho_\xi(t)\,\mathrm{d}t, \quad i = 1 \div N. \tag{3.22}$$

Формулу (3.22) не задовільняє жодна з функцій класу інтегровних функцій, однак у класі узагальнених функцій існує функція, яка задовільняє (3.22) для довільного малого додатнього ε. Це, так звана, δ-функція Дірака. Одним з означень δ-функції Дірака є наступне.

Нехай $\{\varphi_l(x)\}_{l\geqslant 1}$ — функціональна послідовність, елементи якої

$$\varphi_l(x) = \begin{cases} \dfrac{1}{2l}, & |x| < l, \\ 0, & |x| > l. \end{cases} \tag{3.23}$$

Для довільної неперервної функції $g(x)$

$$\int_a^b g(x)\varphi_l(u-x)\,\mathrm{d}x = \int_a^b g(x)\varphi_l(x-u)\,\mathrm{d}x = g(x_0),\ x_0 \in (u-l, u+l), \tag{3.24}$$

а $\forall u \,\overline{\in}\, (a+l, b-l)$

$$\int_a^b g(x)\varphi_l(u-x)\,\mathrm{d}x = \int_a^b g(x)\varphi_l(x-u)\,\mathrm{d}x = 0.$$

То якщо в (3.24) перейти до границі $l \to \infty$, то $\exists \delta(x) = \lim\limits_{l\to\infty} \varphi_l(x)$ в сенсі того, що

$$\int_a^b g(x)\delta(u-x)\,\mathrm{d}x = \int_a^b g(x)\delta(x-u)\,\mathrm{d}x = g(u). \tag{3.25}$$

Таку узагальнену функцію називають функцією Дірака. Використовуючи δ-функцію Дірака, щільність дискретної випадкової величини ξ означають так

$$\rho_\xi(x) = \sum_{i=1}^{N} p_i \delta(x - x_i). \tag{3.26}$$

Справедливо наступне

$$P\{\omega\colon \xi(\omega) = x_k\} = \int\limits_{x_k-\varepsilon}^{x_k-\varepsilon} \rho_\xi(t)\,\mathrm{d}t = p_k.$$

Якщо випадкова величина ξ є дискретно-неперервною (тобто приймає на деякій множині $B \in \mathcal{B}(\mathbb{R})$ дискретні значення, а поза нею є неперервною), то її щільність визначають так:

$$\rho_\xi(x) = g_1(x) + \sum_{i=1}^{N} p_i \delta(x - x_i),$$

де $g_1(x)$ — щільність розподілу (якщо вона існує) неперервної складової випадкової величини $\xi(\omega)$.

Існують випадкові величини, які в жодному інтервалі дійсної осі не є ні дискретними, ні неперервними. Це, так звані, випадкові величини сингулярного типу. Для таких випадкових величин функція розподілу неперервна всюди, але $\frac{\mathrm{d}F}{\mathrm{d}x} = 0$ майже всюди за виключенням множини точок міри нуль. Це означає, що щільність розподілу не існує. Функція розподілу $F(x)$ для таких величин росте лише на множині Лебегової міри нуль, але без стрибків (так звана, крива Кантора $K(x)$).

Справедлива наступна теорема (теорема Лебега). Довільну функцію розподілу $\mathcal{F}$ однозначно можна зобразити у вигляді

$$\mathcal{F}(x) = \sum_{k=1}^{3} b_k F_k(x),$$

де сталі $b_1, b_2, b_3 \geqslant 0$ такі, що

$$b_1 + b_2 + b_3 = 1,$$

а $F_1(x)$ — функція стрибків (дискретний розподіл), $F_2(x)$ — абсолютно-неперервний розподіл і $F_3(x)$ — сингулярна функція розподілу (неперервна функція розподілу без щільності).

§6. Деякі моделі щільностей розподілу

Як було показано в §1 довільна вимірна функція $f(x)$, що задовільняє властивостям:

- $\forall x,\ f(x) \geqslant 0$,
- $\displaystyle\int\limits_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$

може бути щільністю розподілу $\rho_\xi(x)$ абсолютно неперервної випадкової величини ξ.

Тому існує багато модельних щільностей розподілу, серед яких відзначимо наступні.

I. Рівномірний розподіл. Так називають розподіл абсолютно неперервної випадкової величини, щільність розподілу якої є такою:

$$\rho_\xi(x) = \begin{cases} \dfrac{1}{b-a}, & x \in (a,b), \\ 0, & x \overline{\in} (a,b). \end{cases} \tag{3.27}$$

Рівномірний розподіл характерний для випадкових величин, які є грубими (з точністю до цілих поділок шкали) похибками вимірювань. В цьому випадку $a = -\Delta/2$, $b = \Delta/2$, де Δ — ціна поділки шкали.

II. Показниковий розподіл. Так називають розподіл абсолютно неперервної випадкової величини, щільність розподілу якої є

$$\rho_\xi(x) = \lambda \mathrm{e}^{-\lambda x}\theta(x), \quad x \in \mathbb{R}. \tag{3.28}$$

Тут $\lambda > 0$ — стала, яка є параметром розподілу, $\theta(x)$ — функція Хевісайда.

Такий розподіл широко використовується в теорії надійності та теорії обслуговування.

III. Нормальний розподіл (розподіл Гаусса). Такий розподіл визначається щільністю

$$\rho_\xi(x) = \sqrt{\frac{c}{2\pi}} \exp\left\{-\frac{1}{2}c(x-a)^2\right\}, \quad x \in \mathbb{R}, \tag{3.29}$$

де $a \in \mathbb{R}$, $c \in \mathbb{R}_+$ — параметри розподілу.

Позначають його так: $N(c, a)$ ($N(0, 1)$ — *стандартний гауссівський розподіл*). Більш детально цей розподіл, що займає особливе місце серед всіх розподілів, вивчено в розділі VI.

IV. Логарифмічно нормальний (логнормальний) розподіл визначається щільністю

$$\rho_\xi(x) = \sqrt{\frac{c}{2\pi}} \exp\left\{-\frac{1}{2}c(\ln x - a)^2 - a - \frac{1}{2c}\right\}\theta(x), \quad x \in \mathbb{R}, \tag{3.30}$$

де $a \in \mathbb{R}$, $c \in \mathbb{R}_+$ — параметри розподілу.

V. γ-розподіл визначається щільністю

$$\rho_\xi(x) = \frac{k^{p+1}}{\Gamma(\mu + 1)} x^\mu \mathrm{e}^{-kx}\theta(x), \quad x \in \mathbb{R}, \tag{3.31}$$

де $\mu > -1$, $k > 0$ — параметри розподілу,

$$\Gamma(z) = \int\limits_0^\infty \mathrm{e}^{-t} t^{z-1}\,\mathrm{d}t$$

— гамма-функція, відома з курсу математичного аналізу.

Якщо $\mu = n$, $n \in \mathbb{N}$, то γ-розподіл називають *розподілом Ерланга.*

VI. β-розподіл визначається щільністю

$$\rho_\xi(x) = \mathrm{B}(p, q) x^{p-1}(1 - x)^{q-1}\theta(x)\theta(1 - x), \quad x \in \mathbb{R}. \tag{3.32}$$

Тут $p, q > 0$ — параметри розподілу, а

$$\mathrm{B}(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}$$

— відома з курсу математичного аналізу бета-функція.

VII. Розподіл Коші. Щільність розподілу визначається формулою

$$\rho_\xi(x) = \frac{1}{\pi} \frac{a}{x^2 + a^2}, \quad x \in \mathbb{R}, \tag{3.33}$$

де $a > 0$ — параметр розподілу.

VIII. Розподіл Стьюдента визначається щільністю

$$\rho_\xi(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in \mathbb{R}, \quad n = 1, 2, \ldots \tag{3.34}$$

і широко використовується в математичній статистиці.

§7. Ентропія розподілу

Поняття ентропії виникло в теорії інформації як міри невизначеності при побудові двійкового зображення довільного числа

$$Z = a_0 + a_1 2^{-1} + a_2 2^{-2} + \ldots + a_m 2^{-m}. \tag{3.35}$$

Зрозуміло, що таке представлення буде неточним. Якщо за міру неточності подання двійковим числом взяти величину

$$H = m_1 2^{-m_1} + \ldots + m_N 2^{-m_N}, \tag{3.36}$$

де $m_1, \ldots, m_N$ — значення випадкової величини ξ — кількості двійкових знаків в поданні числа Z формулою (3.35), то $p_i = 2^{-m_i}$ є ймовірністю появи значення m_i, тобто $p_i = P\{\omega\colon \xi(\omega) = m_i\}$. Тоді (3.36) можна переписати так

$$H = -\sum_{i=1}^{N} p_i \log_2 p_i. \tag{3.37}$$

Цю величину називають *ентропією*. Одиницею виміру ентропії є двійковий знак — біт. Якщо тепер $A_1, \ldots, A_N$ — повна група несумісних подій з ймовірностями появи $P(A_i) = p_i$, то за міру невизначеності експерименту, в результаті якого появиться одна з них візьмемо величину

$$H = -\sum_{i=1}^{N} p(A_i) \log_2 p(A_i),$$

де

$$\sum_{i=1}^{N} p(A_i) = 1.$$

Оскільки завжди з подібним експериментом можна пов'язати випадкову величину (дискретну) X, закон розподілу якої задано:

$$X = (x_1, \ldots, x_N), \quad P\{x = x_i\} = p_i,$$

то *ентропією випадкової величини (ентропією розподілу)* назвемо величину

$$H[x] = -\sum_{i=1}^{N} p_i \log_2 p_i. \tag{3.38}$$

Зауважимо, що означення $H[x]$ слідує, що ентропія розподілу є неперервною функцією $p_1, \ldots, p_N$; $H[x] \geqslant 0$ (бо $\forall i$, $\log_2 p_i \leqslant 0$).

Якщо x — невипадкова величина, то розглядуваний експеримент є *невипадковим*, $H[x] = 0$. Дійсно, в цьому випадку одна з ймовірностей p_i (p_j) $p_j = 1$, а решта дорівнюють нулю. Так, як $\exists \lim\limits_{p \to 0} p \log_2 p = 0$, то

$$H[x] = -p_j \log_2 p_j = 0.$$

Покажемо, що $H[x]$ досягає максимального значення, якщо $\forall i$, $p_i = \frac{1}{N}$ (рівномірний розподіл).

Використаємо нерівність

$$\forall u \geqslant 0, \ \log_2 u = \frac{\ln u}{\ln 2} \geqslant \frac{1}{\ln 2}\left(1 - \frac{1}{u}\right).$$

Дійсно, нехай $u \geqslant 1$, тоді

$$\ln u = \int\limits_1^u \frac{\mathrm{d}t}{t} \leqslant \frac{1}{u}\int\limits_1^u \mathrm{d}t = \frac{1}{u}(u-1) = 1 - \frac{1}{u}.$$

Якщо $0 \leqslant u \leqslant 1$, то $-\frac{1}{t} \geqslant -\frac{1}{u}$ при $u \leqslant t \leqslant 1$

$$\ln u = -\int_u^1 \frac{\mathrm{d}t}{t} \geqslant -\frac{1}{u}\int_u^1 \mathrm{d}t = 1 - \frac{1}{u}.$$

Тому $\forall u > 0$, $\ln u \leqslant 1 - \frac{1}{u}$.

Нехай тепер $\{q_i\}$ та $\{p_i\}$ $(i = 1 \div N)$ — набори додатніх чисел, таких що

$$\sum_{i=1}^{N} p_i = 1, \quad \sum_{i=1}^{N} q_i = 1.$$

Тоді

$$\sum_{i=1}^{N} p_i \log_2 \frac{p_i}{q_i} \geqslant 0,$$

дійсно,

$$\sum_{i=1}^{N} p_i \log_2 \frac{p_i}{q_i} \geqslant \frac{1}{\ln 2}\sum_{i=1}^{N} p_i \left(1 - \frac{q_i}{p_i}\right) = 0.$$

Знак рівності буде тоді і тільки тоді, коли $\forall i$ $q_i = p_i$. Покладемо, що $\forall i$, $i = 1 \div N$, $q_i = \frac{1}{N}$ і отримуємо, що

$$\begin{aligned}\sum_{i=1}^{N} p_i \log_2 \frac{p_i}{\frac{1}{N}} &= \sum_{i=1}^{N} p_i \log_2 p_i + N \cdot \frac{1}{N}\log_2 N = \\ &= -\left(-\sum_{i=1}^{N} p_i \log_2 p_i - \log_2 N\right) \geqslant 0 \Rightarrow \\ &\Rightarrow -\sum_{i=1}^{N} p_i \log_2 p_i \leqslant \log_2 N.\end{aligned}$$

Знак рівності є *лише* при $p_I = \frac{1}{N}$, $i = 1, 2, \ldots$. Тому

$$H_{\max} = H\left[p_i = \frac{1}{N}\right] = \log_2 N.$$

Отже, ентропія розподілу дискретної випадкової величини досягає свого експериментального значення $\log_2 N$, якщо всі можливі значення випадкової величини x рівноймовірні.

Нехай тепер x — неперервна випадкова величина, задана щільність $\rho_x(x)$. В цьому випадку означимо

$$H[x] = -\int\limits_{-\infty}^{\infty} \rho_x(x) \log_2 \rho_x(x)\,\mathrm{d}x. \tag{3.39}$$

Приклад 3.3. Знайти ентропію показникового розподілу.

Для показникового розподілу

$$\rho_x(x) = \lambda\,\mathrm{e}^{-\lambda x}\theta(x)$$

і тому

$$H[x] = -\int\limits_{0}^{\infty} \lambda\,\mathrm{e}^{-\lambda x}(\log_2 \lambda - \lambda\,x \log_2 \mathrm{e}) = \\ = -\log_2 \lambda + \log_2 \mathrm{e} = \log_2 \frac{\mathrm{e}}{\lambda}.$$

Для неперервних випадкових величин ентропія $H[x]$, взагалі кажучи, можна приймати як додатні, так і від'ємні значення.

Справедлива наступна нерівність: якщо $f(x)$ і $g(x)$ — довільні щільності розподілу, то

$$\int\limits_{-\infty}^{\infty} H(x) \log_2 \frac{f(x)}{g(x)}\,\mathrm{d}x \geqslant 0.$$

Справедлива наступна теорема.

Теорема 3.1. *Якщо ентропія розподілу зі щільністю $f(x)$, що належить до певного класу розподілів C, може бути подана у вигляді*

$$H[x] = -\int_{-\infty}^{\infty} f(x)\log_2 f_0(x)\,\mathrm{d}x,$$

де $f(x)$ — довільна щільність з класу C, то f_0 є єдиним розподілом, що має максимальну ентропію серед всіх розподілів класу C.

Дійсно, нехай Y — випадкова величина зі щільністю $f_Y(y) \in C$. Ентропія

$$H[Y] = -\int_{-\infty}^{\infty} f_Y(y)\log_2 f_Y(y)\,\mathrm{d}y.$$

Тоді

$$H[x] - H[Y] = -\int_{-\infty}^{\infty} f_Y(y)\log_2 f_0(y)\,\mathrm{d}y + \int_{-\infty}^{\infty} f_Y(y)\log_2 f_Y(y)\,\mathrm{d}y =$$

$$= \int_{-\infty}^{\infty} f_Y(y)\log_2 \frac{f_Y(y)}{f_0(y)}\,\mathrm{d}y \geqslant 0.$$

Так як рівність можлива лише тоді і тільки тоді, коли $f_Y(y) = f_0(y)$, то щільність $f_0(y)$ має максимальну ентропію.

Приклад 3.4. Довести, що в класі C n-вимірних розподілів, які визначені в обмеженій області B, максимальну ентропію має рівномірний розподіл.

Нехай x — випадкова величина, розподілена рівномірно в B,

$$\rho_X(x) = \frac{I_B(x)}{\operatorname{meas} B},$$

$I_B(x)$ — індикатор області B. Нехай $f(x)$ — довільна щільність класу C,

$$\int\limits_B f(x)\,\mathrm{d}x = 1 = \int\limits_B \rho_X(x)\,\mathrm{d}x.$$

Тоді

$$\begin{aligned} H[x] = &-\int\limits_B \rho_X(x)\log_2\rho_X(x)\,\mathrm{d}x = \\ &= \log_2 \operatorname{meas} B \cdot \int\limits_B \rho_X(x)\,\mathrm{d}x = \\ &= \log_2 \operatorname{meas} B \cdot \int\limits_B f_X(x)\,\mathrm{d}x = \\ &= -\int\limits_B f_X(x)\log_2\rho_X(x)\,\mathrm{d}x. \end{aligned}$$

За сформульованою вище теоремою щільність $\rho_X(x)$ має максимальну ентропію.

Це значить, що рівномірний розподіл є єдиним розподілом, що має максимальну ентропію серед всіх розподілів неперервних випадкових величин з можливими значеннями, що належать цій же обмеженій області.

ЛЕКЦІЯ 4

Багатовимірні випадкові величини (випадкові вектори) та їх опис

§1. Багатовимірні випадкові величини (випадкові вектори), типи випадкових векторів

Нехай $\{\Omega, \mathfrak{A}, \mathbb{P}\}$ — ймовірнісний простір. Нехай на цьому просторі задано n скалярних випадкових величин $\xi_1(\omega), \ldots, \xi_n(\omega)$.

Означення 1. Довільний впорядкований набір скалярних випадкових величин $(\xi_1(\omega), \ldots, \xi_n(\omega))$ будемо називати n-вимірним випадковим вектором $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$.

Зауважимо, що вектор-функція $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$ здійснює відображення Ω в $\mathbb{R}^n$, тобто

$$\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))\colon \Omega \to \mathbb{R}^n.$$

Так як величини $\xi_1(\omega), \ldots, \xi_n(\omega)$ можна розглядати як координати деякої точки в $\mathbb{R}^n$, то можна стверджувати, що кожній елементарній події $\omega \in \Omega$ ставиться у відповідність вектор $\mathbf{x}$.

Як приклад випадкового вектора, розглянемо задачу про встановлення місцезнаходженння штучного супутника Землі, яке задається набором трьох координат (просторових) та часу:

x_1, x_2, x_3, t. Нехай Δx_i $(i = 1, 2, 3)$ та Δt — випадкові похибки вимірювання. Якщо ω — вимірювання координат супутника, то воно задає чотири величини

$$x_1 = \Delta x_1,\ x_2 = \Delta x_2,\ x_3 = \Delta x_3,\ x_4 = \Delta t$$

— похибка в місцезнаходженні супутника.

Так як і у випадку одновимірних випадкових величин випадкові вектори можуть бути дискретними ($\xi_i(\omega)$ — дискретні), неперервними ($\xi_i(\omega)$ — неперервні), мішані (частина дискретні, частина неперервні).

Як описати випадкові вектори?

§2. Функція розподілу випадкових векторів. Властивості функцій розподілу

Нехай $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$ — випадковий вектор. Нехай випадкова подія

$$A = \{\omega \in \Omega\colon \xi_1(\omega) < x_1, \ldots, \xi_n(\omega) < x_n\} = \\ = \{\omega \in \Omega\colon \xi_i(\omega) < x_i,\ i = 1 \div n\}.$$

Цілком зрозуміло, що $A = \bigcap\limits_{i=1}^{n} B_i$, де $B_i = \{\omega \in \Omega\colon \xi_i(\omega) < x_i\} \in \mathfrak{A}$. Тут $\mathfrak{A}$ — алгебра подій. З властивостей алгебри подій $\mathfrak{A}$ слідує, що

$$A = \bigcap_{i=1}^{n} B_i \in \mathfrak{A}.$$

Тому можна розглядати ймовірність появи події A:

$$P\{\omega \in \Omega\colon \xi_1(\omega) < x_1, \ldots, \xi_n(\omega) < x_n\} = \\ = P\{\omega \in \Omega\colon \xi_i(\omega) < x_i,\ i = 1 \div n\} = \\ = F_{\mathbf{x}}(x_1, \ldots, x_n). \tag{4.1}$$

Означення 2. Функцію $F_{\mathbf{x}}(x_1, \ldots, x_n)$, означену формулою (4.1), будемо називати функцією розподілу випадкового вектора $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$.

Розглянемо властивості цієї функції і покажемо, що вона володіє всіма властивостями функції розподілу одновимірної випадкової величини.

Властивість 1. Згідно з означенням (4.1) $\forall x_i \in \mathbb{R}$ $(i = 1 \div n)$, $F(x_1, \dots, x_n) \in [0, 1]$.

Властивість 2. Функція розподілу $F(x_1, \dots, x_n) \in [0, 1]$ не спадна по кожному зі своїх аргументів, тобто $\forall x_k^{(1)}, x_k^{(2)} \in \mathbb{R}$, $x_k^{(2)} > x_k^{(1)}$

$$F_{\mathbf{x}}(x_1, \dots, x_k^{(2)}, x_{k+1}, \dots, x_n) \geqslant F_{\mathbf{x}}(x_1, \dots, x_k^{(1)}, x_{k+1}, \dots, x_n) \quad (4.2)$$

Доведення. Нехай A_1 та A_2 випадкові події:

$$A_1 = \left\{\omega \in \Omega \colon \xi_k(\omega) < x_k^{(1)}\right\},$$

$$A_2 = \left\{\omega \in \Omega \colon \xi_k(\omega) < x_k^{(2)}\right\}.$$

Якщо

$$B = \{\omega \in \Omega \colon \xi_1(\omega) < x_1, \dots, \xi_{k-1}(\omega) < x_{k-1},$$
$$\xi_{k+1}(\omega) < x_{k+1}, \dots, \xi_n(\omega) < x_n\},$$

то

$$A_2 \cap B = \{A_1 \cap B\} \cup \left\{\left(x_k^{(1)} \leqslant \xi_k(\omega) \leqslant x_k^{(2)}\right) \cap B\right\}.$$

Оскільки $\{A_1 \cap B\} \cap \{(x_k^{(1)} \leqslant \xi_k(\omega) \leqslant x_k^{(2)}) \cap B\} = \varnothing$, то за аксіомою Колмогорова

$$P\{A_2 \cap B\} = P\{A_1 \cap B\} + P\left\{\left(x_k^{(1)} \leqslant \xi_k(\omega) \leqslant x_k^{(2)}\right) \cap B\right\}$$

і тому маємо:

$$P\{A_2 \cap B\} - P\{A_1 \cap B\} = P\left\{\left(x_k^{(1)} \leqslant \xi_k(\omega) \leqslant x_k^{(2)}\right) \cap B\right\} \geqslant 0.$$

Це означає, що

$$F_{\mathbf{x}}(x_1, \dots, x_{k-1}, x_k^{(2)}, x_{k+1}, \dots, x_n) -$$
$$- F_{\mathbf{x}}(x_1, \dots, x_{k-1}, x_k^{(1)}, x_{k+1}, \dots, x_n) \geqslant 0,$$

що і доводить неспадність функції розподілу $F_{\mathbf{x}}(x_1, \ldots, x_n)$ n-вимірного випадкового вектора $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$ по аргументу x_k. Оскільки k — довільне, то ця властивість неспадності $F_{\mathbf{x}}(x_1, \ldots, x_n)$ виконується по кожному зі своїх аргументів. ■

➢ ***Зауваження 4.1.*** *З властивості 2 випливає, що*

$$\begin{aligned} &F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, x_k^{(2)}, x_{k+1}, \ldots, x_n) - \\ &\qquad - F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, x_k^{(1)}, x_{k+1}, \ldots, x_n) = \\ &= P\Big\{\big(x_k^{(1)} \leqslant \xi_k \leqslant x_k^{(2)}\big) \bigcap \big\{\xi_i < x_i,\ i = 1 \div n,\ i \neq k\big\}\Big\}. \end{aligned} \tag{4.3}$$

Введемо різницевий оператор

$$\begin{aligned} \Delta_{a_k, b_k} F_{\mathbf{x}}(x_1, \ldots, x_n) &= F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, b_k, x_{k+1}, \ldots, x_n) - \\ &\quad - F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, a_k, x_{k+1}, \ldots, x_n). \end{aligned} \tag{4.4}$$

Безпосереднім розрахунком можна довести, що

$$\begin{aligned} P(a, b] &= P\big\{(a_1 \leqslant \xi_1 < b_1), (a_2 \leqslant \xi_2 < b_2), \ldots, (a_n \leqslant \xi_n < b_n)\big\} = \\ &= \Delta_{a_1 b_1} \Delta_{a_2 b_2} - \Delta_{a_n b_n} F_{\mathbf{x}}(x_1, \ldots, x_n), \end{aligned} \tag{4.5}$$

де гіперпаралелепіпед

$$(a, b] = (a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n].$$

З (4.5) *випливає, що, взагалі кажучи,*

$$P(a, b] \neq F_{\mathbf{x}}(b, \ldots, b) - F_{\mathbf{x}}(a, \ldots, a).$$

Властивість 3. Для функції розподілу $F_{\mathbf{x}}(x_1, \ldots, x_n)$ випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ справедливі граничні співвідношення 3а)

$$\exists \lim_{\substack{x_i \uparrow \infty \\ i = 1 \div n}} F_{\mathbf{x}}(x_1, \ldots, x_n) = F_{\mathbf{x}}(+\infty, \ldots, +\infty) = 1; \tag{4.6}$$

3b)

$$\exists \lim_{\substack{x_i \to -\infty \\ 1 \leqslant i \leqslant n}} F_{\mathbf{x}}(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = \\ = F_{\mathbf{x}}(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots, x_n) = 0 \qquad (4.7)$$

$\forall \underset{i \neq k}{x_k} \in \mathbb{R}$, які є постійними або змінюються довільним чином.

Доведемо властивість 3. Розглянемо спочатку 3a. Доведемо, що $F_{\mathbf{x}}(+\infty, \ldots, +\infty) = 1$.

Доведення. Нехай випадкова подія

$$A_k = \{\xi_i < k,\ i = 1 \div n\}.$$

Події A_k задовільняють умови

$$\forall k > 1, \quad A_k \subset A_{k+1}; \quad \bigcup_k A_k = \Omega.$$

За аксіомою неперервності

$$\exists \lim_{k \to +\infty} P(A_k) = \lim_{k \to +\infty} F_{\mathbf{x}}(k, \ldots, k) = 1.$$

Нехай тепер $\forall i$, $(1 \leqslant i \leqslant n)$ $x_i \to +\infty$. Зрозуміло, що для всіх k, починаючи з деякого $x_i \geqslant k$, $i = 1 \div n$. За властивістю 2

$$F_{\mathbf{x}}(x_1, \ldots, x_n) \geqslant F_{\mathbf{x}}(k, \ldots, k).$$

Тому

$$\exists \lim_{\substack{x_i \to +\infty \\ i = 1 \div n}} F_{\mathbf{x}}(x_1, \ldots, x_n) = 1.$$

■

Доведемо тепер властивість 3b. Покажемо, що

$$\exists \lim_{\substack{x_k \to -\infty \\ 1 \leqslant k \leqslant n}} F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, x_n) = \\ = F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, -\infty, x_{k+1}, \ldots, x_n) = 0$$

Доведення. За означенням (4.1)

$$F_{\mathbf{x}}(x_1, \ldots, x_n) = P\{\xi_i(\omega) < x_i,\ i = 1 \div n\}.$$

Очевидно, що

$$P\{\xi_i(\omega) < x_i,\ i = 1 \div n\} = P\left\{\bigcap_{i=1}^{n} \{\xi_i(\omega) < x_i\}\right\} \leqslant P\{\xi_i(\omega) < x_i\}$$

(див. теорему множення ймовірностей).

Нехай випадкова подія

$$B_k = \{\xi_i < -k\},$$

тоді справедливі наступні співвідношення:

$$B_{k+1} \subset B_k; \quad \bigcup_k B_k = \varnothing.$$

За теоремою неперервності

$$\exists \lim_{k \to \infty} P(B_k) = 0.$$

Нехай тепер, починаючи з деякого, $x_i < -k$. Внаслідок включення

$$\{\xi_i < x_i\} \subset \{\xi_i < -k\}$$

виконується умова

$$P\{\xi_i < x_i\} \leqslant P\{\xi_i < -k\}.$$

Перейшовши до границі $x_i \to -\infty$ отримуємо, що

$$\exists \lim_{x_i \to -\infty} P\{\xi_i < x_i\} = 0,$$

а тому

$$\exists \lim_{x_i \to -\infty} P\{\xi_i < x_i,\ i = 1 \div n\} = \lim_{x_i \to -\infty} P\{\xi_i < x_i\} = 0.$$

Це означає, що

$$\exists \lim_{\substack{x_k \to -\infty \\ k=1\div n}} F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) =$$
$$= F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, -\infty, x_{k+1}, \ldots, x_n) = 0.$$

■

Ця властивість суттєво відрізняється від властивості функції розподілу скалярної випадкової величини, однак повністю відповідає властивостям алгебри випадкових подій. Дійсно, нехай $A_1, \ldots, A_n \in \Omega$, а $B = \varnothing$. Тоді $\left(\bigcap\limits_{k=1}^{n} A_k\right) \bigcap B = \varnothing$. Властивість 3b є зрозумілою.

Властивість 4 (властивість неперервності). Функція розподілу $F_{\mathbf{x}}(x_1, \ldots, x_n)$ випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ неперервна зліва по всіх своїх аргументах, тобто

$$F_{\mathbf{x}}(x_1, \ldots, x_n) = \lim_{\substack{z_i \uparrow x_i \\ i=1\div n}} F_{\mathbf{x}}(z_1, \ldots, z_n) = F_{\mathbf{x}}(x_1 - 0, \ldots, x_n - 0). \quad (4.8)$$

Доведення. Побудуємо числову послідовність $\{z^{(m)}\}_{m \geqslant 1}$, $z^{(m)} = (z_1^{(m)}, \ldots, z_n^{(m)})$, для якої $z_j^{(m)} \uparrow x_j$, $1 \leqslant j \leqslant n$.

Покажемо, що

$$F_{\mathbf{x}}(z_1^{(m)}, \ldots, z_n^{(m)}) \uparrow F_{\mathbf{x}}(x_1, \ldots, x_n).$$

Нехай S_m — сукупність випадкових подій

$$S_m = \left\{\omega\colon \xi(\omega) < z^{(m)} = \omega\colon \xi_j(\omega) < z_j^{(m)},\ 1 \leqslant j \leqslant n\right\}.$$

Для сукупності подій S_m виконуються (як легко переконатись) умови $\forall m \geqslant 1$, $S_m \subset S_{m+1}$ і

$$\{\omega\colon \xi(\omega) < x\} = \{\omega\colon \xi_j(\omega) < x_j,\ 1 \leqslant j \leqslant n\} = \bigcup_m S_m.$$

За теоремою неперервності

$$\exists \lim_{m \to \infty} P\big(\xi(\omega) < m\big) = P(\xi < x)$$

і тому внаслідок означення $F_{\mathbf{x}}(x_1, \ldots, x_n)$ (формула (4.1)) маємо

$$P\{\omega\colon \xi(\omega) < x\} = F_{\mathbf{x}}(x_1, \ldots, x_n) = \lim_{\substack{z_j \uparrow x_j \\ 1 \leqslant j \leqslant n}} F_{\mathbf{x}}(z_1, \ldots, z_n) = \\ = F_{\mathbf{x}}(x_1 - 0, \ldots, x_n - 0).$$

■

Властивість 5. Нехай $F_{\mathbf{x}}(x_1, \ldots, x_n)$ — функція розподілу випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$. Тоді

$$\exists \lim_{x_k \to +\infty} F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) = \\ = F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, +\infty, x_{k+1}, \ldots, x_n) = \\ = \mathcal{F}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n), \tag{4.9}$$

де $\mathcal{F}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$ — функція розподілу $(n-1)$-вимірного випадкового вектора $\mathbf{x}$.

Доведення. За означенням

$$F_{\mathbf{x}}(x_1, \ldots, x_n) = P\Big\{\omega\colon \bigcap_{i=1}^{n} \{\xi_i(\omega) < x_i\}\Big\}.$$

Так як

$$\lim_{x_k \to +\infty} P\Big\{\omega\colon \bigcap_{i=1}^{n} \{\xi_i(\omega) < x_i\}\Big\} = \\ = P\Big\{\omega\colon \bigcap_{i \neq k}^{n} \{\xi_i(\omega) < x_i\} \cap \{\xi_k(\omega) < +\infty\}\Big\} = \\ = P\Big\{\omega\colon \bigcap_{i=1}^{n-1} \{\xi_i(\omega) < x_i\}\Big\} = \\ = F_{\mathbf{x}}(x_1, \ldots, x_{k-1}, +\infty, x_{k+1}, \ldots, x_n) = \\ = \mathcal{F}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n).$$

■

Надалі функції розподілу

$$\mathcal{F}(x_1, \ldots, x_j, \ldots) = F_{\mathbf{x}}(x_1, \ldots, +\infty, \ldots, x_j, \ldots),$$

які є функціями розподілу випадкового вектора $\mathbf{x}'$, компоненти якого є компонентами випадкового вектора $\mathbf{x}$, що відповідають змінним, які залишились скінченними. Так, наприклад, при $\mathbf{x} = (\xi_1, \xi_2, \xi_3, \xi_4)$ маємо:

$$\mathcal{F}_{\mathbf{x}}(y, u) = F_{\mathbf{x}}(+\infty, y, +\infty, u) = P\{\omega\colon \xi_2(\omega) < y,\, \xi_4(\omega) < u\}.$$

Надалі функції розподілу $\mathcal{F}_{\mathbf{x}}$ будемо називати *частковими (маргінальними)* функціями розподілу. Тоді $F_{\mathbf{x}}$ називають *сукупною* функцією розподілу випадкового вектора $\mathbf{x}$. Знаючи сукупну функцію розподілу $F_{\mathbf{x}}$ можна отримати часткові функції розподілу $\mathcal{F}_{\mathbf{x}}$. Однак знання часткових функцій розподілу не визначає однозначно сукупну функцію розподілу $F_{\mathbf{x}}$.

Справедлива наступна теорема.

Теорема 4.1. *Нехай $F_{\mathbf{x}}(x_1, \ldots, x_n)$ — деяка функція розподілу в $\mathbb{R}^n$. Тоді на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ існує єдина ймовірнісна міра $\mathbb{P}$ така, що*

$$P(a, b] = \Delta_{a_1 b_1} \Delta_{a_2 b_2} \ldots \Delta_{a_n b_n} F_{\mathbf{x}}(x_1, \ldots, x_n).$$

Приклади деяких функцій розподілу випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$.

1. Нехай $F^{(1)}(x_1), \ldots, F^{(n)}(x_n)$ — деякі одновимірні функції розподілу на $\mathbb{R}$. Тоді функція

$$F(x_1, \ldots, x_n) = F^{(1)}(x_1) \cdot F^{(2)}(x_2) \cdot \ldots \cdot F^{(n)}(x_n) \tag{4.10}$$

задовільняє всі властивості 1–4 і як легко перевірити:

$$\Delta_{a_1 b_1} \Delta_{a_2 b_2} \ldots \Delta_{a_n b_n} F_{\mathbf{x}}(x_1, \ldots, x_n) = \prod_{k=1}^{n} \left[F^{(k)}(b_k) - F^{(k)}(a_k) \right] \geqslant 0.$$

Тому $F_{\mathbf{x}}(x_1, \ldots, x_n)$ є функцією розподілу.

Важливим є випадок, коли

$$F^{(k)}(x_k) = \begin{cases} 0, & x_k < 0, \\ x_k, & 0 \leqslant x_k \leqslant 1, \\ 1, & x_k > 1. \end{cases}$$

В цьому випадку для всіх $x_k \in [0,1]$, $k = 1, \ldots, n$

$$F(x_1, \ldots, x_n) = x_1 \cdot \ldots \cdot x_n. \tag{4.11}$$

Відповідну цій функції розподілу ймовірнісну міру $\mathbb{P}$ називають *n-вимірною мірою Лебега на* $[0,1]$.

§3. Дискретні випадкові вектори. Закони розподілу та функції розподілу дискретного випадкового вектора

Нехай $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$ — n-вимірний випадковий вектор.

Означення 3. Вектор $\mathbf{x}$ будемо називати дискретним випадковим вектором, якщо множина його значень $D = D_1 \times D_2 \times \ldots \times D_n$ (D_i — множина значень випадкової величини ξ_i) є скінченною або зліченною.

Означення 4. Законом розподілу n-вимірного випадкового вектора $\mathbf{x} = (\xi_1(\omega), \ldots, \xi_n(\omega))$ будемо називати набір всіх можливих значень $(x_1^i, \ldots, x_n^i)$ випадкового вектора $\mathbf{x}$ та відповідних їм набором ймовірностей

$$P_{i_1,\ldots,i_n} = P\left\{\omega \colon \xi_1 = x_1^{i_1}, \ldots, \xi_n = x_n^{i_n}\right\}. \tag{4.12}$$

При цьому виконується умова нормування

$$\sum_{i_1} \ldots \sum_{i_n} P_{i_1,\ldots,i_n} = 1. \tag{4.13}$$

Табл. 4.1. Закон розподілу двовимірного випадкового вектора $\mathbf{x} = (\xi_1, \xi_2)$.

$\xi_1 \backslash \xi_2$	x_2^1	x_2^2	$\dots$	x_2^m	Σ
x_1^1	p_{11}	p_{12}	$\dots$	p_{1m}	q_1
x_1^2	p_{21}	p_{22}	$\dots$	p_{2m}	q_2
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_1^n	p_{n1}	p_{n2}	$\dots$	p_{nm}	q_m
Σ	p_1	p_2	$\dots$	p_m	1

Зручно ілюструвати закон розподілу випадкового вектора $\mathbf{x}$ для $n = 2$. У випадку $n = 2$, $\mathbf{x} = (\xi_1, \xi_2)$ і закон розподілу випадкового вектора задають у вигляді таблиці 4.1 з двома входами.

В табл. 4.1

$$p_i = \sum_{k=1}^{n} p_{ik}, \quad q_k = \sum_{i=1}^{m} p_{ik}, \tag{4.14}$$

де

$$p_{ik} = \underset{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant k \leqslant m}}{P} \left\{ \omega \colon \xi_1 = x_1^i, \xi_1 = x_2^k \right\}$$

Зауважимо, що

$$\sum_{i=1}^{n} \sum_{k=1}^{m} p_{ik} = 1$$

(умова нормування (4.13)). Ця умова є характерною особливістю закону розподілу (табл. 4.1). З цієї умови слідує, що

$$\sum_{i=1}^{n} p_i = 1, \quad \sum_{k=1}^{m} q_k = 1 \tag{4.15}$$

(довести самостійно).

Знаючи закон розподілу двовимірного випадкового вектора $\mathbf{x} = (\xi_1, \xi_2)$ можна побудувати *маргінальні* розподіли компонент

ξ_1 та ξ_2, тобто знайти ймовірності

$$p_i = P\{\omega\colon \xi_1 = x_i\} = \sum_{k=1}^{m} p_{ik}$$

та

$$q_k = P\{\omega\colon \xi_2 = x_k\} = \sum_{i=1}^{n} p_{ik}$$

(див. табл. 4.2 та 4.3).

Табл. 4.2. Закон розподілу компоненти ξ_1.

$x_1^{(1)}$	$x_1^{(2)}$	$\dots$	$x_1^{(n)}$
p_1	p_2	$\dots$	p_n

де

$$p_i = \sum_{k=1}^{m} p_{ik};$$

Табл. 4.3. Закон розподілу компоненти ξ_2.

$x_2^{(1)}$	$x_2^{(2)}$	$\dots$	$x_2^{(n)}$
q_1	q_2	$\dots$	q_n

де

$$q_k = \sum_{i=1}^{n} p_{ik}.$$

Побудуємо тепер функцію розподілу

$$F_{\mathbf{x}}(x_1, x_2) = \sum_{\substack{(x_1^{(i)})<x_1 \\ (x_2^{(k)})<x_2}} p_{ik}. \tag{4.16}$$

Підсумовування в (4.16) проводимо по всіх значеннях $(x_1^{(i)})$ та $(x_2^{(k)})$, для яких виконуються умови: $x_1^{(i)} < x_1$ та $x_2^{(k)} < x_2$. Проілюструємо викладене.

Приклад 4.1. Задано закон розподілу двовимірного випадкового вектора $\mathbf{x} = (\xi_1, \xi_2)$ (див. табл. 4.4).

Табл. 4.4.

$\xi_1 \backslash \xi_2$	-1	2
-1	0.1	0.15
0	0.05	0.3
2	p_{13}	0.3

Знайти:
1) p_{13};
2) Закон розподілу компоненти ξ_1;
3) Побудувати функцію розподілу;
4) Обчислити $P\{\omega\colon \xi_1 < \xi_2\}$;
1) З умови нормування (4.13) випливає, що

$$p_{13} = 1 - (0.1 + 0.15 + 0.05 + 0.3 + 0.3) = 0.1.$$

2) Побудуємо закон розподілу компоненти ξ_1. Маємо

ξ_1	-1	2
p_i	0.25	0.75

$$p_1 = P\{\omega\colon \xi_1 = -1\} = 0.1 + 0.05 + 0.4 = 0.25,$$
$$p_2 = P\{\omega\colon \xi_1 = 2\} = 0.15 + 0.3 + 0.3 = 0.75 = 1 - 0.25.$$

3) Побудуємо функцію розподілу:

$$F_{\mathbf{x}}(x_1, x_2) = P\{\omega\colon \xi_1 < x_1, \xi_2 < x_2\} =$$
$$= P\{(\xi_1, \xi_2) \in D_{\xi_1\xi_2}\}.$$

Тут

$$D_{\xi_1 \xi_2}\} = \{(y_1, y_2)\colon (y_1, y_2) \in \mathbb{R}^2, y_1 < x_1, y_2 < x_2\}.$$

Зобразимо цю область графічно (див. рис. 4.1). При фіксованих

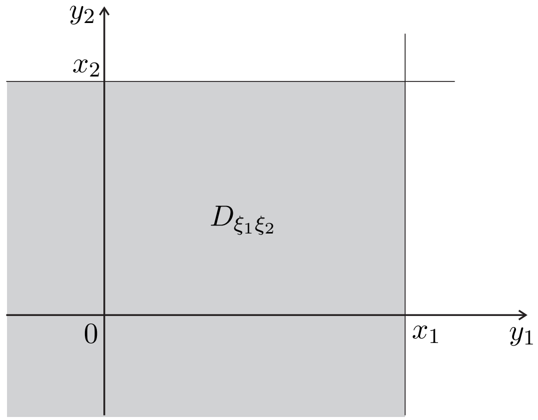

Рис. 4.1. Область $D_{\xi_1\xi_2}$.

x_1 та x_2 значення $F(x_1, x_2)$ дорівнює сумі ймовірностей p_{ik} появи тих значень ξ_1 та ξ_2, які попадають в область $D_{\xi_1\xi_2}$.

Результати розрахунку подано в табл. 4.5.

Табл. 4.5. Функція розподілу $F_{\mathbf{x}}(x_1, x_2)$.

$x_2 \backslash x_1$	$x_1 \leqslant -1$	$-1 < x_1 < 2$	$x_1 > 2$
$x_2 \leqslant -1$	0	0	0
$-1 < x_2 < 0$	0	0.1	0.25
$0 \leqslant x_2 \leqslant 1$	0	0.15	0.7
$x_2 > 1$	0	0.35	1

4) Обчислимо тепер

$$P\{\omega\colon \xi_1 < \xi_2\} = P\{(x_1, x_2)\colon (x_1, x_2) \in \mathbb{R}, x_1 < x_2\}.$$

За теоремою додавання для несумісних подій

$$P\{\xi_1 < \xi_2\} = p_{12} + p_{13} = 0.05 + 0.1 = 0.15.$$

§4. Неперервні випадкові вектори. Щільність розподілу неперервного випадкового вектора

Випадковий n-вимірний вектор $\mathbf{x} = (\xi_1(\omega), \dots, \xi_n(\omega))$ називають *неперервним випадковим вектором*, якщо компоненти $\xi_i(\omega)$, $1 \leqslant i \leqslant n$ є неперервними випадковими величинами, що задані на ймовірнісному просторі $\langle \Omega, \mathfrak{A}, \mathbb{P} \rangle$.

Якщо задана функція розподілу $F(x_1, \dots, x_n)$ (див. означення (4.1)), то для будь-якої борелевої множини B з $\mathbb{R}^n$ можна ввести n-вимірний розподіл ймовірностей

$$\mathbb{P}_\xi(\mathbb{B}) = \{\xi_i(\omega) \in \mathbb{B},\ 1 \leqslant i \leqslant n\}. \tag{4.17}$$

Означення 5. Дійсну невід'ємну інтегровану на $\mathbb{R}^n$ функцію $\rho_{\mathbf{x}}(x_1, \dots, x_n) = \rho_{\mathbf{x}}(\mathbf{x})$ називають *щільністю розподілу неперервного випадкового вектора* $\mathbf{x} = (\xi_1(\omega), \dots, \xi_n(\omega))$ (щільністю спільного розподілу), якщо для кожної борелевої множини $\mathbb{B}$ в $\mathbb{R}^n$

$$\begin{aligned}\mathbb{P}\{\omega\colon \xi_i(\omega) \in \mathbb{B},\ 1 \leqslant i \leqslant n\} &= \int\limits_{\mathbb{B}} \mathrm{d}F_{\mathbf{x}}(\mathbf{x}) = \\ &= \int\limits_{\mathbb{B}} \rho_{\mathbf{x}}(x_1, \dots, x_n)\, \mathrm{d}x_1 \dots \mathrm{d}x_n \end{aligned} \tag{4.18}$$

і

$$\int\limits_{\mathbb{R}^n} \rho_{\mathbf{x}}(x_1, \dots, x_n)\, \mathrm{d}x_1 \dots \mathrm{d}x_n = 1. \tag{4.19}$$

Умову (4.19) називають *умовою нормування.*

➢ ***Зауваження 4.2.*** *Взагалі кажучи, інтеграл в означенні (4.18) є інтегралом Лебега. Однак надалі такі інтеграли будемо розуміти як інтеграли Рімана.*

Поклавши в означенні (4.18)

$$\mathbb{B} = (-\infty, x_1] \times \ldots \times (-\infty, x_n],$$

то використовуючи (4.1), отримуємо, що

$$F_{\mathbf{x}}(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} \rho_{\xi}(t_1, \ldots, t_n)\, \mathrm{d}t_1 \ldots \mathrm{d}t_n. \qquad (4.20)$$

Якщо в (4.20) кратний інтеграл можна звести до повторного (для цього достатньо неперервності $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$ в $\mathbb{R}^n$), то скрізь в точках неперервності

$$\rho_{\mathbf{x}}(x_1, \ldots, x_n) = \frac{\partial^n F_{\mathbf{x}}(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n}. \qquad (4.21)$$

Введена формулами (4.20), чи (4.21) щільність спільного розподілу неперервного вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ має такі властивості:
1) щільність спільного розподілу $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$ невід'ємна в $\mathbb{R}^n$, тобто

$$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n \qquad \rho_{\mathbf{x}}(x_1, \ldots, x_n) \geqslant 0;$$

2) щільність спільного розподілу $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$ нормована на одиницю, тобто

$$\exists \int\limits_{\mathbb{R}^n} \rho_{\mathbf{x}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n = 1.$$

Ця властивість дозволяє стверджувати, що щільність спільного розподілу $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$ належить до класу інтегровних в $\mathbb{R}^n$ функцій. Справедливе і обернене твердження: довільна невід'ємна інтегровна на $\mathbb{R}^n$ функція, що задовільняє умову нормування (4.14) є щільністю сукупного розподілу деякого n-вимірного випадкового вектора.

Якщо задано сукупну щільність розподілу $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$, то, скориставшись означенням (4.9) часткових функцій розподілу, можна побудувати *часткові щільності розподілу.* Дійсно, нехай $\mathbf{X} = (\xi_1, \ldots, \xi_n)$ та $\mathbf{Y} = (\xi_1, \ldots, \xi_k)$ випадкові вектори $(k < n)$ з тими ж компонентами, але різної вимірності. Покладемо в означенні (4.20) $x_i = +\infty$, $k < i \leqslant n$ і отримуємо

$$\mathbb{P}\{\omega\colon \xi_j(\omega) < x_j,\, 1 \leqslant j \leqslant k\} = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_k} g_{\mathbf{Y}}(t_1, \ldots, t_k)\, \mathrm{d}t_1 \ldots \mathrm{d}t_k, \tag{4.22}$$

де

$$g_{\mathbf{Y}}(t_1, \ldots, t_k) = \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} \rho_{\mathbf{X}}(x_1, \ldots, x_k, t_{k+1}, \ldots, t_n)\, \mathrm{d}t_{k+1} \ldots \mathrm{d}t_n. \tag{4.23}$$

Це означає, що $g_{\mathbf{Y}}(x_1, \ldots, x_k)$ є сукупною щільністю розподілу вектора $\mathbf{Y} = (\xi_1, \ldots, \xi_k)$, вимірність якого дорівнює k. Тому формула (4.23) дозволяє знайти щільності сукупного розподілу довільної кількості компонент $\xi_1, \ldots, \xi_k$, $k < n$ випадкового вектора $\mathbf{X} = (\xi_1, \ldots . \xi_n)$.

Зокрема, щільність розподілу компоненти ξ_i є

$$\begin{aligned}\rho_{\xi_i}(x) = \int\limits_{-\infty}^{+\infty} \mathrm{d}t_1 \int\limits_{-\infty}^{+\infty} \mathrm{d}t_2 \ldots \int\limits_{-\infty}^{+\infty} \mathrm{d}t_{i-1} \int\limits_{-\infty}^{+\infty} \mathrm{d}t_{i+1} \ldots \times \\ \times \int\limits_{-\infty}^{+\infty} \mathrm{d}t_n \rho_{\mathbf{X}}(t_1, \ldots, t_{i-1}, x, t_{i+1}, \ldots, t_n), \quad i = 1 \div n.\end{aligned} \tag{4.24}$$

Викладене вище означає, що знаючи сукупну щільність розподілу $\rho_{\mathbf{X}}(x_1, \ldots, x_n)$ випадкового вектора $\mathbf{X} = (\xi_1, \ldots, \xi_n)$ можна обчислити щільності розподілу компонент (формула (4.24)). Обернене твердження, взагалі кажучи, невірне.

На практиці найбільш вживаними моделями щільностей розподілу випадкових векторів $\mathbf{X} = (\xi_1, \ldots, \xi_n)$ є такі.

- **Рівномірний розподіл.**

Нехай V — вимірна обмежена область в $\mathbb{R}^n$. Випадкова величина $\mathbf{X} = (\xi_1, \dots, \xi_n)$ рівномірно розподілена в області V, якщо

$$\rho_{\mathbf{X}}(\xi_1, \dots, \xi_n) = \begin{cases} \dfrac{1}{\operatorname{meas} V}, & (x_1, \dots, x_n) \in V, \\ 0, & (x_1, \dots, x_n) \overline{\in} V. \end{cases}$$

• **Гаусів (нормальний) розподіл в $\mathbb{R}^n$.**

Нехай $\mathbf{A} = \|a_{ij}\|$ — симетрична додатньо визначена матриця, вимірності $\tau(\mathbf{A}) = n \times n$, і нехай

$$\mathbf{a} = \left\| \begin{matrix} a_1 \\ \vdots \\ a_n \end{matrix} \right\|, \quad \text{а} \quad \mathbf{X} = \left\| \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \right\|,$$

$(x_1, \dots, x_n) \in \mathbb{R}^n$.

Щільність гаусового (нормального) закону розподілу задається виразом:

$$\rho_{\mathbf{X}}(x_1, \dots, x_n) = \sqrt{\frac{\det \mathbf{A}}{(2\pi)^n}} \exp\left\{ -\frac{1}{2} (\mathbf{X} - \mathbf{a})^{\mathrm{T}} \mathbf{A} (\mathbf{X} - \mathbf{a}) \right\}. \tag{4.25}$$

У випадку $n = 2$ щільність (4.25) можна привести шляхом алгебраїчних перетворень до

$$N^2(a_1, a_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \times \\ \times \exp\left\{ -\frac{1}{2(1 - \rho^2)} \left[\frac{(x_1 - a_1)^2}{\sigma_1^2} + \frac{(x_2 - a_2)^2}{\sigma_2^2} - \right.\right. \\ \left.\left. - 2\rho \frac{(x_1 - a_1)(x_2 - a_2)}{\sigma_1\sigma_2} \right] \right\}. \tag{4.26}$$

Параметри розподілу: $a_1, a_2 \in \mathbb{R}$, $\sigma_1 > 0$, $\sigma_2 > 0$, $|\rho| < 1$.

Приклад 4.2. Випадкова величина $\boldsymbol{\xi} = (\xi_1, \xi_2)$ рівномірно розподілена в одиничному крузі з центром в початку координат. Знайти:

а) функцію розподілу $F_\xi(x_1, x_2)$;

б) щільність розподілу компоненти ξ_1.

Так як випадкова величина розподілена рівномірно в крузі радіуса $R = 1$ з центром в початку координат, то

$$\rho_{\boldsymbol{\xi}}(x_1, x_2) = \begin{cases} \dfrac{1}{\pi}, & x_1^2 + x_2^2 \leqslant 1, \\ 0, & x_1^2 + x_2^2 > 1. \end{cases}$$

Обчислимо за означенням (4.24)

$$\rho_{\xi_1}(x_1) = \int\limits_{-\infty}^{+\infty} \mathrm{d}x_2\, \rho_{\boldsymbol{\xi}}(x_1, x_2) =$$
$$= \begin{cases} \dfrac{2}{\pi}\sqrt{1 - x_1^2}, & |x_1| \leqslant 1, \\ \quad 0, & |x_1| > 1. \end{cases}$$

Скориставшись означенням (4.20) функції розподілу $F_{\boldsymbol{\xi}}(x_1, x_2)$ маємо:

$$F_{\boldsymbol{\xi}}(x_1, x_2) = \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{x_2} \rho_{\boldsymbol{\xi}}(t_1, t_2)\, \mathrm{d}t_1\, \mathrm{d}t_2 =$$
$$= \mathbb{P}\{\omega \in \Omega \colon \xi_1(\omega) < x_1,\ \xi_2(\omega) < x_2\}.$$

Розрахунок інтегралу можна провести безпосередньо (проробити самостійно). Ми ж скористаємося поняттям геометричної ймовірності.

Нехай B — область в $\mathbb{R}^2$:

$$B = (-\infty, x_1] \times (-\infty, x_2].$$

Нехай D — круг одиничного радіуса з центром в початку координат:

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \colon x_1^2 + x_2^2 \leqslant 1\}.$$

Розглянемо область

$$T \stackrel{\text{def}}{=} B \cap D.$$

Щільність сумісного розподілу випадкового вектора $\boldsymbol{\xi} = (\xi_1, \xi_2)$:

$$\rho_{\boldsymbol{\xi}}(x_1, x_2) = \begin{cases} \dfrac{1}{\mu(D)}, & (x_1, x_2) \in D, \\ 0, & (x_1, x_2) \overline{\in} D. \end{cases}$$

Тут $\mu(D)$ — міра області D (для нашого випадку $\rho(D) = \pi$). Тому

$$\mathbb{P}\{\omega \in \Omega \colon \xi_1(\omega) < x_1,\, \xi_2(\omega) < x_2\} = \iint\limits_{\mathfrak{F}} \rho_{\boldsymbol{\xi}}(t_1, t_2)\, \mathrm{d}t_1\, \mathrm{d}t_2 =$$

$$= \frac{1}{\mu(D)} \iint\limits_{T} \mathrm{d}t_1\, \mathrm{d}t_2 = \frac{\mu(B \cap D)}{\mu(D)}.$$

Отже, отримано загальний вираз

$$F_{\boldsymbol{\xi}}(x_1, x_2) = \frac{\mu(B \cap D)}{\mu(D)}.$$

Тому для розрахунку слід обчислити $\mu(B \cap D) = \mu(T)$.

а) Нехай $B \cap D = \varnothing$, $\mu(\varnothing) = 0$, $F_{\boldsymbol{\xi}}(x_1, x_2) = 0$. Ця ситуація зображена на рис. 4.2, тобто якщо $x_1 \in (-\infty, -1] \cup x_2 \in \mathbb{R}$ або $x_2 \in (-\infty, -1] \cup x_1 \in \mathbb{R}$, то $F_{\boldsymbol{\xi}}(x_1, x_2) = 0$.

б) Нехай $B \cap D = D$. Це відповідає випадку, коли $D \subset B$. $F_{\boldsymbol{\xi}}(x_1, x_2) = 1$. Ця ситуація зображена на рис. 4.3 і є випадком, коли $x_1 \in (1, +\infty] \cup x_2 \in (1, +\infty)$.

в) Нехай $B \cap D \neq \varnothing$ та $B \cap D \neq D$. Можлива ситуація зображена на рис. 4.4. Заштрихована область є $B \cap D$. Обчисли-

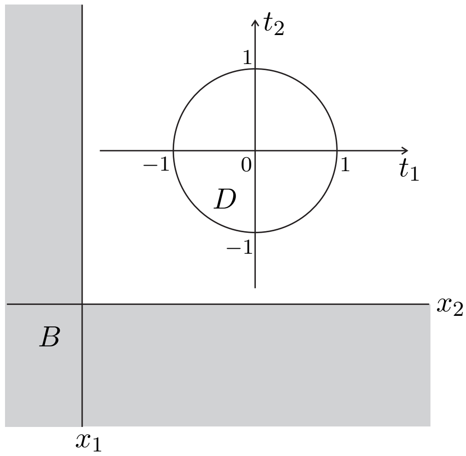

Рис. 4.2.

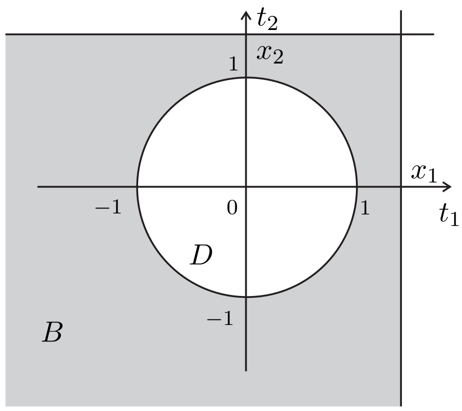

Рис. 4.3.

мо $\mu(B \cap D)$

$$\mu(B \cap D) = \iint\limits_{\mathfrak{F}\cap D} \mathrm{d}t_1\,\mathrm{d}t_2 = \int\limits_{x_1^0}^{x_1}\mathrm{d}t_1 \int\limits_{-\sqrt{1-t_1^2}}^{x_2}\mathrm{d}t_2 =$$

$$= \int\limits_{x_1^0}^{x_1}\mathrm{d}t\left(x_2 + \sqrt{1-t^2}\right) =$$

$$= x_2(x_1 - x_1^0) + \left(\frac{t}{2}(1-t^2)^{1/2} + \frac{1}{2}\arcsin t\right)\Big|_{x_1^0}^{x_1} =$$

$$= x_2(x_1 - x_1^0) + \frac{x_1}{2}(1-x_1^2)^{1/2} + \frac{1}{2}\arcsin x_1 -$$

$$- \frac{x_1^0}{2}\left(1-(x_1^0)^2\right)^{1/2} - \frac{1}{2}\arcsin x_1^0.$$

Оскільки (як видно з рис. 4.3) $x_1^0 = -\sqrt{1-x_2^2}$ і тому

$$F_{\boldsymbol{\xi}}(x_1, x_2) = x_1x_2 + \frac{x_1}{2}(1-x_1^2)^{1/2} + \frac{x_2}{2}(1-x_2^2)^{1/2} +$$

$$+ \frac{1}{2}\arcsin x_1 + \frac{1}{2}\arcsin x_2,\ x_1 \in (-1,1] \cup x_2 \in (-1,1].$$

§5. Залежність компонент випадкового вектора. Умовні закони розподілу

Якщо відома функція розподілу випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$, то, як показано в §2 цього розділу, можна знайти функції розподілу компонент

$$F_{\xi_i}(x_i) = F_{\mathbf{x}}(+\infty, +\infty, \ldots, x_i, +\infty, \ldots, +\infty).$$

Аналогічно можна визначити щільності розподілу компонент, якщо відома щільність розподілу випадкового вектора, а саме:

$$\rho_{\xi_i}(x_i) = \int\limits_{-\infty}^{+\infty}\ldots\int\limits_{-\infty}^{+\infty}\rho_{\mathbf{x}}(t_1, \ldots, t_{i-1}, x_i, t_{i+1}, \ldots, t_n)\,\mathrm{d}t_1\ldots\,\mathrm{d}t_n.$$

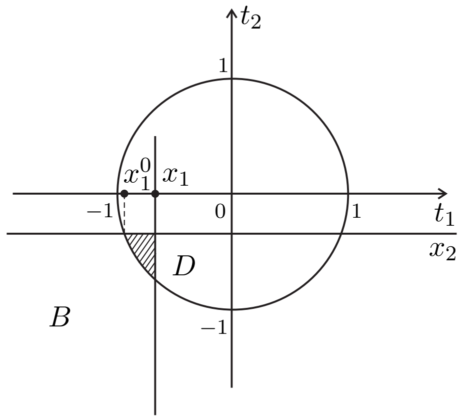

Рис. 4.4.

Поставимо зворотню задачу: відомі щільності (чи функції) розподілу компонент. Чи можна встановити функції (чи щільності) розподілу випадкового вектора? Це не завжди можливо, оскільки необхідно вияснити як взаємозв'язані між собою компоненти випадкового вектора.

Нехай $\langle \Omega, \mathfrak{A}, \mathbb{P} \rangle$ — ймовірнісний простір. Будемо говорити, що випадкові величини $\xi_1(\omega), \dots, \xi_n(\omega)$ (компоненти випадкового вектора) задані на ймовірнісному просторі *незалежні* (чи незалежні в сукупності), якщо

$$\mathbb{P}\{\omega \in \Omega \colon \xi_i(\omega) \in B_i,\, 1 \leqslant i \leqslant n\} = \\ = \mathbb{P}\{\omega \in \Omega \colon \xi_1(\omega) \in B_1\} \dots \mathbb{P}\{\omega \in \Omega \colon \xi_n(\omega) \in B_n\}, \quad \forall B_i \in \mathbb{B}(\mathbb{R}). \tag{4.27}$$

Якщо

$$F_{\mathbf{x}}(x_1, \dots, x_n) = \mathbb{P}\{\omega \in \Omega \colon \xi_i(\omega) < x_i,\, 1 \leqslant i \leqslant n\}$$

— функція розподілу n-вимірного випадкового вектора $\mathbf{x} = (\xi_1, \dots, x_n)$, а F_{ξ_i}, $i = 1, \dots, n$ — функція розподілу випадкової величини ξ_i (i-ої компоненти випадкового вектора $\mathbf{x}$), то справедлива наступна теорема.

Теорема 4.2. *Для того, щоб випадкові величини* $\xi_i(\omega)$, $1 \leqslant i \leqslant n$ *були незалежні, необхідно і достатньо, щоб для всіх* $(x_1, \dots, x_n) \in \mathbb{R}^n$

$$F_{\mathbf{x}}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n). \tag{4.28}$$

Доведення. *Необхідність.*

Покажемо, що якщо $\xi_i(\omega)$, $1 \leqslant i \leqslant n$ незалежні, то

$$F_{\mathbf{x}}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i).$$

Це випливає з означення (4.27), так як

$$F_{\mathbf{x}}(x_1, \dots, x_n) = \mathbb{P}\{\omega \in \Omega \colon \xi_i(\omega) < x_i,\, 1 \leqslant i \leqslant n\},$$

а $F_{\xi_i}(x_i) = \mathbb{P}\{\omega \in \Omega \colon \xi_i(\omega) < x_i\}$, то необхідність є очевидною.

Достатність.

Нехай $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{b} = (b_1, \dots, b_n)$. Покладемо

$$\mathbb{P}_{\mathbf{x}}(\mathbf{a}, \mathbf{b}] = P\{\omega \in \Omega \colon a_i < \xi_i(\omega) \leqslant b_i,\, 1 \leqslant i \leqslant n\},$$

$$\mathbb{P}_{\xi_i}(a_i, b_i] = P\{\omega \in \Omega \colon a_i < x_i \leqslant b_i, \}, \quad i = 1 \div n.$$

Тоді (див. формулу (4.5))

$$\mathbb{P}_{\mathbf{x}}(\mathbf{a}, \mathbf{b}] = \Delta_{a_1 b_1} \dots \Delta_{a_n b_n} F_{\mathbf{x}}(x_1, \dots, x_n).$$

За умовою теореми:

$$F_{\mathbf{x}}(x_1, \dots, x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i),$$

тому

$$P_x(\mathbf{a}, \mathbf{b}] = \prod_{i=1}^{n} \left[F_{\xi_i}(b_i) - F_{\xi_i}(b_i)\right].$$

Використавши (4.5), маємо

$$P_{\xi_i}(a_i, b_i] = F_{\xi_i}(b_i) - F_{\xi_i}(a_i), \quad i = 1 \div n.$$

Тому

$$P_{\mathbf{x}}(\mathbf{a}, \mathbf{b}] = \prod_{i=1}^{n} P_{\xi_i}(a_i, b_i]. \tag{4.29}$$

Достатність доведена для випадку відрізків $(a_i, b_i]$. Використовуючи злічену аддитивність ймовірнісної міри, можна показати, що (4.29) справедливе і для довільних борелівських множин $B_i \in \in \mathbb{B}(\mathbb{R})$. Тому справедливе означення (4.27), а значить достатність доведено. ■

Нехай існує сукупна щільність розподілу $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$ випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ означена формулою (4.21). Тоді дану теорему можна переформулювати для щільностей розподілу. Справедлива наступна теорема.

Теорема 4.3. *Для того, щоб абсолютно неперервні випадкові величини $\xi_i(\omega)$, $i = 1 \div n$ були незалежні, необхідно і достатньо, щоб для всіх $(x_1, \ldots, x_n) \in \mathbb{R}^n$*

$$\rho_{\mathbf{x}}(x_1, \ldots, x_n) = \rho_{\xi_1}(x_1) \ldots \rho_{\xi_n}(x_n). \tag{4.30}$$

Якщо компоненти випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ залежні, то вводять поняття *умовних* законів розподілу.

Для спрощення розглянемо випадок $n = 2$.

I. Нехай $\mathbf{x} = (\xi_1, \xi_2)$ — двовимірний випадковий вектор і нехай його компоненти ξ_1 і ξ_2 є дискретними випадковими величинами.

Умовним законом розподілу компоненти ξ_1 будемо називати розподіл ймовірностей $\mathbb{P}\{\omega \in \Omega\colon (\xi = x_1)|(\xi_2 = x_2)\}$, який легко отримати, знаючи розподіл ймовірностей $\mathbb{P}\{\omega \in \Omega\colon \xi = x_1, \xi_2 = = x_2\}$. З теореми множення для залежних подій маємо

$$\mathbb{P}\{\omega \in \Omega\colon \xi_1 = x_1 | \xi_2 = x_2\} = \mathbb{P}\{\omega \in \Omega\colon \xi_1 = x_1, \xi_2 = x_2\} = \\ = \frac{P\{(\xi_1 = x_1) \cap (\xi_2 = x_2)\}}{P\{\xi_2 = x_2\}}. \tag{4.31}$$

Розподіл ймовірності $P\{\xi_2 = x_2\}$ відомий з сукупного закону розподілу $P\{\xi_1 = x_1, \xi_2 = x_2\} = P\{\omega \in \Omega\colon \{\xi_1 = x_1\} \cap \{\xi_2 = x_2\}\}$. Аналогічно умовний закон розподілу компоненти ξ_2 є таким

$$\mathbb{P}\{\omega \in \Omega\colon (\xi_2 = x_2)|(\xi_1 = x_1)\} = \mathbb{P}\{\omega \in \Omega\colon \xi_2 = x_2|\xi_1 = x_1\} = \\ = \frac{P\{\xi_1 = x_1, \xi_2 = x_2\}}{P\{\xi_1 = x_1\}}. \tag{4.32}$$

Зауважимо, що умовні закони розподілу є нормованими, тобто виконується умова

$$\sum_i P\left\{(\xi_1 = x_1^{(i)})|(\xi_2 = x_2^{(i)})\right\} = 1. \tag{4.33}$$

Для незалежних випадкових величини умовні закони розподілу співпадають із законами розподілу компонент (переконайтесь самостійно).

II. Розглянемо тепер випадок абсолютно неперервних випадкових компонент ξ_1 та ξ_2 випадкового вектора $\mathbf{x} = (\xi_1, \xi_2)$.

Нехай $\rho_{\mathbf{x}}(\xi_1, \xi_2)$ — сукупна щільність розподілу вектора $\mathbf{x}$.

Позначимо через $\rho_{\xi_1|\xi_2}(x_1|x_2)$ умовну щільність розподілу компоненти ξ_1 за умови, що компонента ξ_2 набула значення $\xi_2 = x_2$. Для обчислення $\rho_{\xi_1|\xi_2}(x_1|x_2)$ поступимо так. Знайдемо ймовірність $\mathbb{P}\{\omega \in \Omega\colon \{\xi_1 < x_1\} \cap \{x_2 \leqslant \xi_2 < x_2 + \Delta x\}\}$. За теоремою множення

$$\mathbb{P}\{\omega \in \Omega\colon \{\xi_1 < x_1\} \cap \{x_2 \leqslant \xi_2 < x_2 + \Delta x\}\} = \\ = \mathbb{P}\{\omega \in \Omega\colon x_2 \leqslant \xi_2 < x_2 + \Delta x\} \times \\ \times \mathbb{P}\{\omega \in \Omega\colon (\xi_1 < x_1)|(x_2 \leqslant \xi_2 < x_2 + \Delta x)\}.$$

Якщо $\mathbb{P}\{\omega \in \Omega\colon x_2 \leqslant \xi_2 < x_2 + \Delta x\} \neq 0$, то

$$\mathbb{P}\{\omega \in \Omega\colon (\xi_1 < x_1)|(x_2 \leqslant \xi_2 < x_2 + \Delta x)\} = \\ = \frac{\mathbb{P}\{\omega \in \Omega\colon \{\xi_1 < x_1\} \cap \{x_2 \leqslant \xi_2 < x_2 + \Delta x\}\}}{\mathbb{P}\{\omega \in \Omega\colon (x_2 \leqslant \xi_2 < x_2 + \Delta x)\}} \tag{4.34}$$

за умови, що $\mathbb{P}\{\omega \in \Omega\colon (x_2 \leqslant \xi_2 < x_2 + \Delta x)\} \neq 0$.

Так як

$$\mathbb{P}\{\omega \in \Omega\colon \{\xi_1 < x_1\} \cap \{x_2 < \xi_2 \leqslant x_2 + \Delta x\}\} = \\ = \int\limits_{-\infty}^{x_1} \mathrm{d}t_1 \int\limits_{x_2}^{x_2+\Delta x} \mathrm{d}t_2\, \rho_{\mathbf{x}}(t_1, t_2), \tag{4.35}$$

а

$$\mathbb{P}\{\omega \in \Omega\colon (x_2 \leqslant \xi_2 < x_2 + \Delta x\}\} = \\ = \int\limits_{x_2}^{x_2+\Delta x} \rho_{\xi_2}\, \mathrm{d}t_2 \equiv \int\limits_{x_2}^{x_2+\Delta x} \mathrm{d}t_2 \int\limits_{-\infty}^{+\infty} \mathrm{d}t_1\, \rho_{\mathbf{x}}(t_1, t_2), \tag{4.36}$$

тому

$$\mathbb{P}\{\omega \in \Omega\colon (\xi_1 < x_1)|(x_2 \leqslant \xi_2 < x_2 + \Delta x)\} = \\ = \frac{\int\limits_{-\infty}^{x_1} \mathrm{d}t_1 \int\limits_{x_2}^{x_2+\Delta x} \rho_{\mathbf{x}}(t_1, t_2)\, \mathrm{d}t_2}{\int\limits_{x_2}^{x_2+\Delta x} \mathrm{d}t_2 \int\limits_{-\infty}^{+\infty} \mathrm{d}t_1\, \rho_{\mathbf{x}}(t_1, t_2)} = \frac{\frac{1}{\Delta x} \int\limits_{-\infty}^{x_1} \mathrm{d}t_1 \int\limits_{x_2}^{x_2+\Delta x} \rho_{\mathbf{x}}(t_1, t_2)\, \mathrm{d}t_2}{\frac{1}{\Delta x} \int\limits_{-\infty}^{+\infty} \mathrm{d}t_1 \int\limits_{x_2}^{x_2+\Delta x} \rho_{\mathbf{x}}(t_1, t_2)\, \mathrm{d}t_2}. \tag{4.37}$$

Перейдемо в (4.37) до границі коли $\Delta x \to 0$, тобто знайдемо

$$\lim_{\Delta x \to 0} \mathbb{P}\{\omega \in \Omega\colon (\xi_1 < x_1)|(x_2 \leqslant \xi_2 < x_2 + \Delta x)\} = \\ = \mathbb{P}\{\omega \in \Omega\colon (\xi_1 < x_1)|(\xi_2 = x_2)\} = \\ = \frac{\int\limits_{-\infty}^{x_1} \mathrm{d}t_1\, \rho_{\mathbf{x}}(t_1, x_2)}{\int\limits_{-\infty}^{+\infty} \mathrm{d}t_1\, \rho_{\mathbf{x}}(t_1, x_2)} = \frac{\int\limits_{-\infty}^{x_1} \rho_{\mathbf{x}}(t, x_2)\, \mathrm{d}t}{\rho_{\xi_2}(x_2)} = \int\limits_{-\infty}^{x_1} \frac{\rho_{\mathbf{x}}(t, x_2)}{\rho_{\xi_2}(x_2)}\, \mathrm{d}t.$$

Означивши умовну щільність розподілу $\rho_{\xi_1|\xi_2}(x_1|x_2)$ через співвідношення

$$\mathbb{P}\{\omega \in \Omega\colon (\xi_1 < x_1)|(\xi_2 = x_2)\} = \int\limits_{-\infty}^{x_1} \rho_{\xi_1|\xi_2}(t|x_2)\,\mathrm{d}t, \tag{4.38}$$

отримуємо, що

$$\rho_{\xi_1|\xi_2}(x_1|x_2) = \frac{\rho_{\mathbf{x}}(x_1, x_2)}{\rho_{\xi_2}(x_2)}, \quad \rho_{\xi_2}(x_2) \neq 0 \tag{4.39}$$

— умовна щільність розподілу компоненти ξ_1 за умови, що $\xi_2 = x_2$.

Аналогічно

$$\rho_{\xi_2|\xi_1}(x_2|x_1) = \frac{\rho_{\mathbf{x}}(x_1, x_2)}{\rho_{\xi_1}(x_1)}, \quad \rho_{\xi_1}(x_1) \neq 0 \tag{4.40}$$

— умовна щільність розподілу компоненти ξ_2 за умови, що $\xi_1 = x_1$.

Використавши (4.39) та (4.40), маємо:

$$\rho_{\mathbf{x}}(\xi_1, \xi_2) = \rho_{\xi_1|\xi_2}(x_1|x_2)\rho_{\xi_2}(x_2) = \rho_{\xi_2|\xi_1}(x_2|x_1)\rho_{\xi_1}(x_1). \tag{4.41}$$

Якщо $\rho_{\xi_2}(x_2) = 0$ (чи $\rho_{\xi_1}(x_1) = 0$), то покладають, що $\rho_{\xi_1|\xi_2}(x_1|x_2) = 0$ (чи $\rho_{\xi_2|\xi_1}(x_2|x_1) = 0$).

Якщо випадкові величини $\xi_1(\omega)$ та $\xi_2(\omega)$ — незалежні, то

$$\rho_{\xi_1|\xi_2}(x_1|x_2) = \rho_{\xi_1}(x_1),$$

$$\rho_{\xi_2|\xi_1}(x_2|x_1) = \rho_{\xi_2}(x_2),$$

що випливає з теореми незалежності компонент та формули (4.41).

Розглянемо приклад.

Приклад 4.3. Нехай (ξ_1, ξ_2) — двовимірна випадкова величина, компоненти якої ξ_1 та ξ_2 незалежні випадкові величини, розподілені за нормальними законами $N(a_1, \sigma_1)$ та $N(a_2, \sigma_2)$. Знайти сукупну щільність розподілу величини (ξ_1, ξ_2).

За формулою (4.30)

$$\rho_{(\xi_1\xi_2)}(x_1, x_2) = \rho_{\xi_1}(x_1)\rho_{\xi_2}(x_2).$$

Так як

$$\rho_{\xi_1}(x_1) = N(a_1, \sigma_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(x_1 - a_1)^2}{2\sigma_1^2}\right\},$$

$$\rho_{\xi_2}(x_2) = N(a_2, \sigma_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(x_2 - a_2)^2}{2\sigma_2^2}\right\},$$

то

$$\rho_{(\xi_1\xi_2)}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(x_1 - a_1)^2}{2\sigma_1^2} - \frac{(x_2 - a_2)^2}{2\sigma_2^2}\right\}.$$

Порівнюючи цей вираз з виразом (4.26), бачимо, що для незалежних випадкових величин ξ_1 і ξ_2 параметр $\rho = 0$. Питання про роль параметра ρ буде розглянуте в розділі 7.

Приклад 4.4. Нехай випадкова величина ξ розподілена за експоненціальним законом, тобто

$$\rho_\xi(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Знайти умовний розподіл $\mathbb{P}\{(\xi - a) \leqslant x | \xi \geqslant a\}$, $a > 0$.

Маємо

$$\mathbb{P}\{(\xi - a) \leqslant x | \xi \geqslant a\} = \frac{\mathbb{P}\{a \leqslant \xi \leqslant a + x\}}{\mathbb{P}\{\xi \geqslant a\}},$$

$$\mathbb{P}\{a \leqslant \xi \leqslant a + x\} = F_\xi(a + x) - F_\xi(a) + \mathbb{P}\{\xi = a\},$$

а

$$\mathbb{P}\{\xi \geqslant a\} = 1 - F_\xi(a) + \mathbb{P}\{\xi = a\},$$

де

$$F_\xi(x) = \begin{cases} 1 - \exp(-\lambda x), & x \geqslant 0, \\ 0, & x < 0 \end{cases}$$

— функція розподілу експоненціально розподіленої випадкової величини x і $\mathbb{P}\{\xi = a\} = 0$, маємо

$$\begin{aligned}\mathbb{P}\{(\xi - a) < x|\xi \geqslant a\} &= \frac{[1 - \exp(-\lambda(a + x))] - [1 - \exp(-\lambda a)]}{1 - [1 - \exp(-\lambda a)]} = \\ &= \frac{\exp(-\lambda a)[1 - \exp(-\lambda x)]}{\exp(-\lambda a)} = \\ &= 1 - \exp(-\lambda x) = F_\xi(x).\end{aligned}$$

Це означає, що

$$\mathbb{P}\{(\xi - a) < x|\xi \geqslant a\} = \mathbb{P}\{\xi \leqslant x\}.$$

Ця властивість експоненціального закону є характеристичною, оскільки не існує інших розподілів зі щільностями, для яких

$$\mathbb{P}\{(\xi - a) \leqslant x|\xi \geqslant a\} = \mathbb{P}\{\xi \leqslant x\}, \quad a > 0, \quad x \in \mathbb{R}_+.$$

ЛЕКЦІЯ 5

Функції випадкових величин

§1. Функція скалярної випадкової величини. Ряд та функція розподілу функції випадкової величини. Щільність розподілу функції випадкової величини

Значна кількість практичних задач вимагає дослідження випадкових величин, які є функціями скалярних (одновимірних) випадкових величин $\xi(\omega)$ з відомими для них законами, чи функціями розподілу.

Наприклад, побудувати розподіл випадкового значення амплітуди A радіосигналу, яка змінюється за законом $A(t) = a(1 + \cos t)$ у випадковий момент часу τ, що рівномірно розподілений в інтервалі $(-\pi, \pi)$. Для вияснення цих питань введемо поняття функції випадкової величини ξ[1]. Нехай $\langle \Omega, \mathfrak{A}, \mathbb{P} \rangle$ — ймовірнісний простір, на якому визначено випадкову величину $\xi(\omega)$. Нехай $g(x)$ — борелівська функція; $g\colon \mathbb{R}^1 \to \mathbb{R}^1$.

Означення 1. Складену функцію $\eta = g(\xi(\omega))$ називають *функцією випадкової величини* ξ.

Справедлива наступна лема.

[1] Як і раніше розглядаємо лише дійсні випадкові величини.

➭ ***Лема 5.1.*** *Нехай* $g = g(x)$ — *борелівська функція, а* $\xi(\omega)$ — *випадкова величина. Тоді функція* $\eta = g(\xi(\omega))$ *є випадковою величиною.*

Доведення. Доведення випливає з того, що для всіх борелівських підмножин $B \in \mathcal{B}(\mathbb{R})$

$$\{\omega\colon \eta(\omega) \in B\} = \{\omega\colon g(\xi(\omega)) \in B\} = \\ = \{\omega\colon \xi(\omega) \in g^{-1}(B)\} \in \mathfrak{A},$$

оскільки $g^{-1}(B) \in \mathcal{B}(\mathbb{R})$. ■

Доведена лема стверджує, що якщо ξ є випадковою величиною, то такі функції, як наприклад, ξ^n, $|\xi|$, $\sin\xi$ також є випадковими величинами, оскільки функції x^n, $|x|$, $\sin x$ є борелівськими.

Нехай задано функцію розподілу $F_\xi(x)$ випадкової величини $\xi = \xi(\omega)$, а $y = g(x)$ — деяка борелівська функція. Розглянемо основні способи знаходження функцій розподілу $F_\eta(y)$ функції $\eta = g(\xi)$.

I. Дискретні випадкові величини.

Нехай $\xi(\omega)$ — дискретна випадкова величина із заданим рядом розподілу $p_i = \mathbb{P}\{\xi = x_i\}$, $i \in \mathbb{N}$, $\sum\limits_i p_i = 1$.

Розглянемо функцію $\eta = g(\xi)$, де $g\colon \mathbb{R}^1 \to \mathbb{R}^1$ — строго монотонна борелівська функція. Зрозуміло, що множина значень функції $\eta = g(\xi)$ буде дискретною множиною значень виду

$$y_j = g(\xi_j) = g(x_j), \quad j \in \mathbb{N}. \tag{5.1}$$

Знайдемо розподіл ймовірностей цих значень. Так як події $\{\eta = = y_j = g(x_j)\}$ та $\xi = x_j$ еквівалентні, то

$$\begin{aligned} &q_j = \mathbb{P}\{\eta = y_j\} = \mathbb{P}\{\xi = x_j\}, \\ &y_j = g(x_j), \quad j \in \mathbb{N}. \end{aligned} \tag{5.2}$$

Причому $\sum\limits_j q_j = 1$.

Отож, якщо відображення $g\colon \mathbb{R}^1 \to \mathbb{R}^1$ є однозначним (функція $g(x)$ — строго монотонна) і борелівським, то ряд розподілу функції $\eta = g(\xi)$ дискретної випадкової величини $\xi(\omega)$ задається системою рівностей (5.1) та (5.2):

$$y_i = g(x_i), \quad P\{\eta = y_i\} = P\{\xi = x_i\}, \quad i \in \mathbb{N}. \tag{5.3}$$

Якщо ж відображення g є немонотонним (функція $g(x)$ — многозначна, тобто одному значенню функції $\eta = y_j$ відповідає певна підмножина значень $\xi = x_j^{(k)}$, $k = 1 \div s$, $y_j = \left\{\begin{array}{c} g(x_j^{(1)}) \\ \vdots \\ g(x_j^{(s)}) \end{array}\right.$), то справедлива еквівалентність наступних подій

$$\{\omega \in \Omega\colon \eta(\omega) = y_j\} = \bigcup_{k=1}^{s} \left\{\omega \in \Omega\colon \xi(\omega) = x_j^{(k)}\right\},$$

причому події $\{\xi = x_j^{(k)}\}$ попарно несумісні для різних k. За теоремою додавання

$$q_j = \mathbb{P}\{\eta = y_j\} = \sum_{k\colon g(x_j^{(k)}) = y_j} \mathbb{P}\{\xi = x_j^{(k)}\}. \tag{5.4}$$

Підсумовування в (5.3) поширюється на ті значення $x_j^{(k)}$, для яких $y_j = g(x_j^{(k)})$. Якщо відомий закон розподілу функції $y_j = g(\xi)$, то як розглянуто в розділі IV, легко розраховується функція розподілу $F_\eta(y)$, де $y = g(x)$.

Приклад 5.1. Дискретна випадкова величина ξ задана рядом розподілу

ξ	-1	0	1	2
p	0.2	0.1	0.3	0.4

Побудувати ряд розподілу та функцію розподілу $F_\eta(y)$ функції $\eta = \xi^2$.

Обчислимо значення випадкової величини η:

$$y_1 = x_1^2 = 1, \quad y_2 = x_2^2 = 0,$$

$$y_3 = x_3^2 = 1 = y_1, \quad y_2 = x_4^2 = 4.$$

Та як одному значенню $y_1 = 1$ відповідає два значення $x_1 = -1$ та $x_3 = 1$, то

$$\begin{aligned} q_1 &= P\{\eta = y_1\} = 0.2 + 0.3 = 0.5, \\ q_2 &= P\{\eta = y_2\} = 0.1, \\ q_3 &= P\{\eta = y_4\} = 0.4 \end{aligned}$$

і ряд розподілу функції $\eta = \xi^2$ є таким

η	0	1	4
q	0.1	0.5	0.4

Функція розподілу $F_\eta(y) = \mathbb{P}\{\eta < y\}$ (як показано в розділі IV) дорівнює

$$F_\eta(y) = \begin{cases} 0, & y \in (-\infty, 0), \\ 0.1, & y \in [0, 1), \\ 0.6, & y \in [1, 4), \\ 1, & y \in [4, +\infty). \end{cases}$$

II. Абсолютно неперервні випадкові величини.

Нехай ξ — абсолютно неперервна випадкова величина, для якої задано функцію розподілу $F_\xi(x)$, а щільність розподілу $\rho_\xi(x)$ величини ξ. Якщо $y = g(x)$ є визначеною, неперервно диференційовною функцією і строго монотонною (тобто $g'(x)$ має сталий знак) в області значень $I = (a, b)$ випадкової величини ξ, то позначивши $I_y = (-\infty, y)$, знаходимо для випадкової величини $\eta = g(x)$:

$$\begin{aligned} F_\eta(y) &= \mathbb{P}\{\eta < y\} = \mathbb{P}\{g(\xi) \in I_y\} = \\ &= \mathbb{P}\left\{\xi \in g^{-1}(I_y)\right\} = \int\limits_{g^{-1}(I_y)} F_\xi(\mathrm{d}x). \end{aligned} \tag{5.5}$$

Інтеграл в (5.5) є інтегралом Лебега-Стільтьєса.

Якщо відома щільність $\rho_\xi(x)$, то для (5.5) маємо:

$$F_\eta(y) = \mathbb{P}\{\eta < y\} = \int\limits_{g^{-1}(I_y)} \rho_\xi \, \mathrm{d}x. \tag{5.6}$$

Нехай $f_\eta(y)$ — щільність розподілу випадкової величини $\eta = g(\xi)$. Тоді з врахуванням (5.6) отримуємо рівність:

$$F_\eta(y) = \int\limits_{I(y)} f_\eta(t) \, \mathrm{d}t = \int\limits_{g^{-1}(I_y)} \rho_\xi \, \mathrm{d}x. \tag{5.7}$$

Зробимо в другому інтегралі заміну змінних $x = g^{-1}(y)$. Це можливо, бо для функції $y = g(x)$ за рахунок сталості знаку похідної $g'(x)$ існує обернена функція $x = g^{-1}(y)$. Тому (5.6) можна переписати так:

$$F_\eta(y) = \int\limits_{I_y} f_\eta(y) \, \mathrm{d}y = \int\limits_{I_y} \rho\big(g^{-1}(y)\big) |J(y)| \, \mathrm{d}y, \tag{5.8}$$

де

$$|J(y)| = \left| \frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \right|.$$

З (5.8) випливає, що

$$f_\eta(y) = \rho\big(g^{-1}(y)\big) \left| \frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \right|, \quad y \in (-\infty, y]. \tag{5.9}$$

Вираз (5.9) задає щільність розподілу $f_\eta(y)$ функції $\eta = g(\xi)$ випадкової величини ξ через щільність розподілу $\rho_\xi(x)$ аргументу ξ.

Якщо функція $y = g(x)$ є строго зростаючою ($g'(x) > 0$), то

$$f_\eta(y) = \rho\big(g^{-1}(y)\big) \frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}. \tag{5.10}$$

Якщо функція $y = g(x)$ є строго спадною, то

$$f_\eta(y) = \rho\big(g^{-1}(y)\big) \left(-\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \right). \tag{5.11}$$

Приклад 5.2. Нехай випадкова величина ξ розподілена зі щільністю $\rho_\xi(x)$, $x \in \mathbb{R}$. Знайти щільність розподілу функції $\eta = \xi^k$, де k — непарне натуральне число.

Для непарних k існує єдина обернена функція $x = y^{1/k}$. Так як $\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} = \frac{1}{k}\, y^{1/k-1}$, то згідно з (5.9)

$$f_\eta(y) = \frac{1}{k}\, \rho_\xi\big(y^{1/k}\big) \left| y^{1/k-1} \right| .$$

Якщо функція $y = g(x)$ не є строго монотонною, то вираз (5.9) для $f_\eta(y)$, $\eta = g(\xi)$, взагалі кажучи, незастосовний. Однак для багатьох практичних задач повністю достатнім є таке узагальнення (5.9).

Нехай функція $y = g(x)$ визначена на деякій множині $\Delta = \bigcup\limits_{k=1}^{n} \Delta_k$, $\Delta_k = [a_k, b_k]$, причому на кожному відкритому інтервалі $I_k = (a_k, b_k)$ $g(x)$ є неперервно диференційовною функцією ($\forall x \in I_k$, $\exists g'(x) \neq 0$). Позначимо через $g_k^{-1}(y)$ обернену до $g(x)$ функцію, $x \in I_k$. Тоді формулу (5.9) легко узагальнити, а саме:

$$f_\eta(y) = \sum_{k=1}^{n} \rho_\xi\big(g_k^{-1}(y)\big) \left| \frac{\mathrm{d}g_k^{-1}(y)}{\mathrm{d}y} \right| I_{D_k}. \tag{5.12}$$

В (5.12) I_{D_k} — індикатор множини D_k — області визначення функції $g_k^{-1}(y)$. Таке узагальнення легко утримується з (5.9). Для цього область інтегрування $g^{-1}(I_y)$ слід розбити на області монотонності функції $y = g(x)$, і для кожної з цих областей скористатись формулою (5.9), а далі використати властивість адитивності інтегралу відносно області інтегрування.

Приклад 5.3. Знайти щільність розподілу випадкової величини $\eta = \cos\xi$, якщо випадкова величина ξ розподілена рівномірно в інтервалі $(-\pi, \pi)$.

Зобразимо графік функції $y = \cos x$, $x \in (-\pi, \pi)$. В даному випадку для кожного значення y, $|y| < 1$ обернена функція

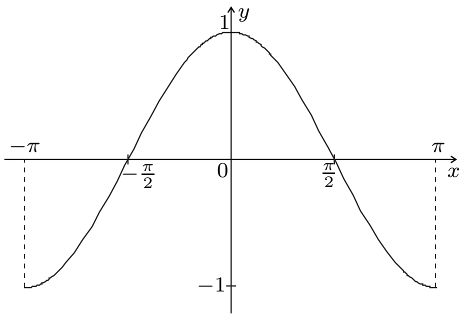

Рис. 5.1.

$x = g^{-1}(y)$ має дві вітки: $g_1^{-1}(y) = \arccos y$, $g_2^{-1}(y) = -\arccos y$. При чому $g_1^{-1}(y)$ визначена для $y \in [0,1)$, а $g_2^{-1}(y)$ визначена для $y \in (-1, 0)$. Так як

$$\rho_\xi(x) = \begin{cases} \dfrac{1}{2\pi}, & x \in (-\pi, \pi), \\ 0, & x \overline{\in} (-\pi, \pi), \end{cases}$$

то використовуючи (5.12), отримуємо, що

$$f_\eta(y) = \begin{cases} \dfrac{1}{2\pi}\left(\dfrac{1}{\sqrt{1-y^2}} + \dfrac{1}{\sqrt{1-y^2}}\right), & |y| < 1, \\ 0, & y \in (-\infty, -1) \cup (1, +\infty) \end{cases}$$

або

$$f_\eta(y) = \frac{1}{\sqrt{1-y^2}}\,\theta(1-|y|),$$

де $\theta(x)$ — функція Хевісайда.

§2. Функції випадкового вектора. Закони розподілу та щільності розподілу функцій випадкового вектора

Узагальнимо розглянуті в попередньому параграфі поняття на випадок функцій випадкових векторів. Нехай $\langle \Omega, \mathfrak{A}, \mathbb{P} \rangle$

— ймовірнісний простір, на якому визначено систему випадкових величин $\xi_1(\omega), \dots, \xi_n(\omega)$[2] або задано випадковий вектор $\mathbf{x} = (\xi_1, \dots, \xi_n)$. Нехай $y_k = g_k(x_1, \dots, x_n)$, $1 \leqslant k \leqslant m$ сукупність борелівських функцій $g_k \colon \mathbb{R}^n \to \mathbb{R}^m$. Справедливе наступне означення.

Означення 2. Складену функцію $Y = (\eta_1, \dots, \eta_m)$, де $\eta_k = g_k(\xi_1, \dots, \xi_n)$, $1 \leqslant k \leqslant m$, називають *функцією випадкового вектора* $\mathbf{x}$.

Надалі для спрощення запису будемо використовувати позначення

$$\mathbf{Y} = G(\mathbf{X}), \tag{5.13}$$

в якому

$$\mathbf{Y} = (\eta_1, \dots, \eta_m), \quad \mathbf{X} = (\xi_1, \dots, \xi_n),$$

$$G = (g_1(\xi_1, \dots, \xi_n), \dots, g_m(\xi_1, \dots, \xi_n)).$$

Означення 2 забезпечує виконання леми про випадковість функції випадкової величини[3]. Це дозволяє стверджувати, що компоненти

$$\eta_k = g_k(\xi_1, \dots, \xi_n), \quad 1 \leqslant k \leqslant m$$

вектора $\mathbf{Y}$ є випадковими величинами, а значить вектор $\mathbf{Y} = (\eta_1, \dots, \eta_m)$ є випадковим вектором.

Розглянемо тепер основні способи побудови закону розподілу $\mathbb{P}\{\eta_j = y_j, 1 \leqslant j \leqslant m\}$ (чи сукупної щільності розподілу $f_{\mathbf{Y}}(y_1, \dots, y_m)$) випадкового вектора $\mathbf{Y} = G(\mathbf{X})$ (функції) за заданими законами розподілу $\mathbb{P}\{\xi_i = x_i, 1 \leqslant i \leqslant n\}$ (чи щільностями розподілу $\rho_{\mathbf{X}}(x_1, \dots, x_n)$) вектора $\mathbf{X}$ (аргумента).

[2] Для простоти будемо вважати, що $\forall i$, $1 \leqslant i \leqslant n$ $\xi_i(\omega)$ — дійсні.

[3] При доведенні цієї леми не використовувалась розмірність множини B.

I. Сумісний закон розподілу функції дискретного вектора. Нехай $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ є дмскретним випадковим вектором із заданим сукупним розподілом

$$p_{1,\ldots,n} = \mathbb{P}\{\omega \in \Omega\colon \xi_1 = x_1, \ldots, \xi_n = x_n\}$$

і нехай $y_k = g_k(x_1, \ldots, x_n)$, $1 \leqslant k \leqslant m$ — сукупність борелівських функцій, кожна з яких здійснює відображення $\mathbb{R}^n \to \mathbb{R}^1$.

Розглянемо випадковий вектор $\mathbf{Y} = (\eta_1, \ldots, \eta_m)$, компоненти якого $\eta_k = g_k(\xi_1, \ldots, \xi_n)$, $1 \leqslant k \leqslant m$. Так як компоненти ξ_j, $1 \leqslant j \leqslant n$ вектора $\mathbf{X}$ приймають дискретні значення $\xi_j = x_j$, то компоненти η_k вектора $\mathbf{Y}$ теж будуть дискретними і прийматимуть значення $y_k = g_k(x_1, \ldots, x_n)$, $1 \leqslant k \leqslant m$.

Сукупний розподіл ймовірностей вектора $\mathbf{Y} = G(\mathbf{X})$

$$q_{1,\ldots,m} = \mathbb{P}\{\omega \in \Omega\colon \eta_1 = y_1, \ldots, \eta_n = y_n\},$$

де $y_k = g_k(x_1, \ldots, x_n) \in \mathbb{R}$, $1 \leqslant k \leqslant m$ отримуємо з наступних міркувань.

Розглянемо систему рівнянь

$$y_k = g_k(1_1, \ldots, x_n), \quad 1 \leqslant k \leqslant m. \tag{5.14}$$

Нехай Δ — множина розв'язків цієї системи $\Delta = (x_1^0, \ldots, x_n^0)$, тобто виконуеться умова

$$y_k^0 = g_k(x_1^0, \ldots, x_n^0), \quad 1 \leqslant k \leqslant m. \tag{5.15}$$

Позначимо через B множину значень $y_k^0 = g_k(x_1^0, \ldots, x_n^0)$, $1 \leqslant k \leqslant m$. Справедлива наступна еквівалентність подій:

$$\begin{aligned}&\{\omega \in \Omega\colon \eta_k(\omega) = y_k^0, 1 \leqslant k \leqslant m, y_k^0 \in B\} = \\ &\quad = \{\omega \in \Omega\colon \xi_i(\omega) = x_i^0, 1 \leqslant i \leqslant n, x_i^0 \in \Delta = G^{-1}(B)\},\end{aligned} \tag{5.16}$$

з якої випливає,що

$$\begin{aligned}&\mathbb{P}\{\omega \in \Omega\colon \eta_k(\omega) = y_k^0, 1 \leqslant k \leqslant m, y_k^0 \in B\} = \\ &\quad = \mathbb{P}\{\omega \in \Omega\colon \xi_i(\omega) = x_i^0, 1 \leqslant i \leqslant n, x_i^0 \in \Delta = G^{-1}(B)\},\end{aligned} \tag{5.17}$$

Використання (5.16) та (5.17) дозволяє виразити сукупний закон розподілу дискретного випадкового вектора $\mathbf{Y} = G(\mathbf{X})$ через сукупний закон розподілу аргумента $\mathbf{X}$ наступним співвідношенням:

$$q_{1,\dots,m} = \sum_{\{x_i\colon x_i \in \Delta\}} p_{1,\dots,n}. \tag{5.18}$$

Знаючи сукупний закон розподілу, можна розрахувати функцію розподілу $F_{\mathbf{Y}}(y_1, \dots, y_m)$.

З іншого боку, скориставшись означенням сукупної функції розподілу

$$F_{\mathbf{Y}}(y_1, \dots, y_m) = \mathbb{P}\{\omega \in \Omega\colon \eta_1(\omega) < y_1, \dots, \eta_m(\omega) < y_m\},$$

отримуємо, що

$$F_{\mathbf{Y}}(y_1, \dots, y_m) = \sum_{\{x_i\colon x_i \in D\}} p_{1,\dots,n}. \tag{5.19}$$

Підсумовування в (5.19) ведеться по тих значеннях x_i, що належать області D — множині розв'язків системи нерівностей

$$y_k > g_k(x_1, \dots, x_n), \quad 1 \leqslant k \leqslant m.$$

Розглянемо наступний приклад.

Приклад 5.4. Задано сукупний закон розподілу ймовірностей дискретного випадкового вектора $\mathbf{x} = (\xi_1, \xi_2)$

$\xi_1 \backslash \xi_2$	-2	-1	0	1
-1	0.01	0.02	0.05	0.03
0	0.03	0.24	0.15	0.06
1	0.06	0.09	0.16	0.10

Знайти закон розподілу випадкової величини $Y = \xi_2 - |\xi_1|$.

Для кожної пари $(x_1^{(i)}, x_2^{(j)})$ значень випадкових величин ξ_1 та ξ_2 обчислюємо значення $y_k = y(x_1^{(i)}, x_2^{(j)}) = x_2^i - |x_1^i|$ знаходимо ймовірність $p_{ij} = \mathbb{P}\{\xi_1 = x_1^i, \xi_2 = x_2^j\}$ та складемо таблицю

$(x_1^{(i)}, x_2^{(j)})$	$(-1,-2)$	$(-1,-1)$	$(-1,0)$	$(-1,1)$	$(0,-2)$	$(0,-1)$
y_k	-3	-2	-1	0	-2	-1
p_{ij}	0.01	0.02	0.05	0.03	0.03	0.24

$(x_1^{(i)}, x_2^{(j)})$	$(0,0)$	$(0,1)$	$(1,-2)$	$(1,-1)$	$(1,0)$	$(1,1)$
y_k	0	1	-3	-2	-1	0
p_{ij}	0.15	0.06	0.06	0.09	0.16	0.1

Впорядкувавши значення y_k (в послідовності зростання), знаходимо ймовірності, що їм відповідають і отримуємо закон розподілу $Y = \xi_2 - |\xi_1|$

y_k	-3	-2	-1	0	1
$P(Y = y_k)$	0.07	0.14	0.45	0.28	0.06

Легко переконатися, що $\sum\limits_{k=1}^{5} P(Y = y_k) = 1$.

II. Щільність розподілу функції неперервного випадкового вектора. Нехай $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ — неперервний випадковий вектор із заданою сукупною щільністю $\rho_{\mathbf{X}}(x_1, \ldots, x_n)$. Нехай $G = (g_1(x_1, \ldots, x_n), \ldots, g_k(x_1, \ldots, x_n))$ — сукупність борелівських функцій, кожна з яких здійснює відображення $\mathbb{R}^n \to \mathbb{R}^1$.

Випадковий вектор $\mathbf{Y} = G(\mathbf{X})$, компонента $\eta_1, \ldots, \eta_m$ якого є функціями випадкових величин $\xi_1, \ldots, \xi_n$:

$$\eta_k = g_k(\xi_1, \ldots, \xi_n), \quad 1 \leqslant k \leqslant m$$

має сукупну щільність розподілу $f_{\mathbf{Y}}(y_1, \ldots, y_n)$.

Будемо вважати, що вимірності векторів $\mathbf{X}$ та $\mathbf{Y}$ співпадають $(m = n)$. Припустимо, що система рівнянь $y_k = g_k(z_1, \ldots, x_n)$, $1 \leqslant$ $\leqslant k \leqslant n$ для довільних $y_k \in \mathbb{R}$ має єдиний розв'язок $(x_1^0, \ldots, x_n^0)$ в області можливих значень координат випадкового вектора $\mathbb{X}$ (в області де $\rho_{\mathbf{X}}(x_1, \ldots, x_m) \neq 0$). Для цього необхідно і достатньо, щоб якобіан переходу від змінних $(x_1, \ldots, x_n)$ до змінних $(y_1, \ldots, y_n)$ зберігав знак в області можливих значень координат вектора $\mathbf{X}$, тобто $\forall (x_1, \ldots, x_n) \in \mathbb{R}^n$:

$$J(y) = \frac{D(x_1, \ldots, x_n)}{D(y_1, \ldots, y_n)} = \frac{D\big(g_1^{-1}(y_1, \ldots, y_n), \ldots, g_n^{-1}(y_1, \ldots, y_n)\big)}{D(y_1, \ldots, y_n)} =$$
$$= \begin{vmatrix} \dfrac{\partial g_1^{-1}(y_1, \ldots, y_n)}{\partial y_1} & \ldots & \dfrac{\partial g_1^{-1}(y_1, \ldots, y_n)}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_n^{-1}(y_1, \ldots, y_n)}{\partial y_1} & \ldots & \dfrac{\partial g_n^{-1}(y_1, \ldots, y_n)}{\partial y_n} \end{vmatrix} \neq 0. \tag{5.20}$$

Нехай $B = \bigcup\limits_{k=1}^{n} B_k$, де $B_k = (-\infty, y_k)$. Тоді згідно з означенням (4.20)

$$F_{\mathbf{Y}}(y_1, \ldots, y_n) = \mathbb{P}\{\omega \in \Omega \colon \eta_k(\omega) < y_k, 1 \leqslant k \leqslant n\} =$$
$$= \int\limits_B f_{\mathbf{Y}}(y_1, \ldots, y_n)\, \mathrm{d}y_1 \ldots \mathrm{d}y_n. \tag{5.21}$$

З іншого боку

$$\mathbb{P}\{\eta_1 < y_1, \ldots, \eta_n < y_n\} = \int\limits_{\Delta_B} \rho_{\mathbf{X}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n, \tag{5.22}$$

де $\Delta_B = G^{-1}(B)$ — прообраз множини B;

$$\Delta_B = ((x_1, \ldots, x_n) \in \mathbb{R}^n \colon g_k(x_1, \ldots, x_n) < y_k, 1 \leqslant k \leqslant n.)$$

Тому можна записати наступну рівність

$$\int\limits_B f_{\mathbf{Y}}(y_1, \ldots, y_n)\, \mathrm{d}y_1 \ldots \mathrm{d}y_n = \int\limits_{\Delta_B} \rho_{\mathbf{X}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n. \tag{5.23}$$

Цю рівність називають *принципом відповідності ймовірностей.*

Зробимо в останньому інтегралі заміну змінних

$$x_i = g_i^{-1}(y_1, \ldots, y_n), \quad 1 \leqslant i \leqslant n.$$

Маємо

$$F_{\mathbf{Y}}(y_1,\dots,y_n)=\int\limits_B f_{\mathbf{Y}}(y_1,\dots,y_n)\,\mathrm{d}y_1\dots\,\mathrm{d}y_n=$$
$$=\int\limits_{\Delta_B}\rho_{\mathbf{X}}\big(g_1^{-1}(y_1,\dots,y_n),\dots,g_n^{-1}(y_1,\dots,y_n)\big)\times$$
$$\times\,|J(y_1,\dots,y_n)|\,\mathrm{d}y_1\dots\,\mathrm{d}y_n, \tag{5.24}$$

а сукупна щільність вектора $\mathbf{Y}=G(\mathbf{X})$ є такою:

$$f_{\mathbf{Y}}(y_1,\dots,y_n)=\rho_{\mathbf{X}}\big(g_1^{-1}(y_1,\dots,y_n),\dots,g_n^{-1}(y_1,\dots,y_n)\big)\times$$
$$\times\,|J(y_1,\dots,y_n)|. \tag{5.25}$$

Приклад 5.5. Знайти розподіл полярних координат точки (R,θ) на площині, якщо відома щільність розподілу $\rho_{\mathbf{X}}(x_1,x_2)$ декартових координат (x_1,x_2) точки.

Нехай $R=\sqrt{x_1^2+x_2^2}=g_1(x_1,x_2)$, а $\theta=\operatorname{arctg}\frac{x_2}{x_1}=g_2(x_1,x_2)$, причому величини кута θ визначаються (як відомо з курсу алгебри) лише знаками величин x_1 та x_2, а саме:

$$x_1>0,\,x_2>0\Rightarrow\theta\in\left[0,\frac{\pi}{2}\right],$$
$$x_1<0,\,x_2>0\Rightarrow\theta\in\left(\frac{\pi}{2},\pi\right],$$
$$x_1<0,\,x_2<0\Rightarrow\theta\in\left(\pi,\frac{3\pi}{2}\right],$$
$$x_1>0,\,x_2<0\Rightarrow\theta\in\left(\frac{3\pi}{2},\pi2\right).$$

Система рівнянь

$$\begin{cases} R=g_1(x_1,x_2)=\sqrt{x_1^2+x_2^2},\\ \theta=g_2(x_1,x_2)=\operatorname{arctg}\dfrac{x_2}{x_1}\end{cases}$$

має єдиний розв'язок

$$x_1 = g_1^{-1}(R, \theta) = R\cos\theta,$$
$$x_2 = g_2^{-1}(R, \theta) = R\sin\theta$$

для всіх $R \in [0, +\infty)$, $\theta \in [0, 2\pi)$.

Так як

$$J(R, \theta) = \begin{vmatrix} \cos\theta & -R\sin\theta \\ \sin\theta & R\cos\theta \end{vmatrix} = R,$$

то згідно з (5.18)

$$F_{\mathbf{Y}}(R, \theta) = R \cdot \rho_{\mathbf{X}}(R\sin\theta, R\cos\theta).$$

Якщо система рівнянь

$$y_j = g_j(x_1, \ldots, x_n), \quad 1 \leqslant j \leqslant n, \tag{5.26}$$

$y_j \in \mathbb{R}^1$, має множину розв'язків $\Delta(y)$ $(\Delta(y) = ((x_1, \ldots, x_n) \in \in \mathbb{R}\colon x_i = g_i^{-1}(y_1,]\ldots, y_n), 1 \leqslant i \leqslant n))$ і якщо ця множина допускає розбиття на множини $I_k(y)$ — множини єдиності розв'язку системи (5.26)

$$\Delta(Y) = \bigcup_k I_k(Y),$$

то формула (5.18) має узагальнення, а саме:

$$f_{\mathbf{Y}}(y_1, \ldots, y_n) = \sum_{I_k(Y) \in \Delta(Y)} \rho_{\mathbf{X}}\big(g_1^{-1}(y_1, \ldots, y_n), \ldots, g_n^{-1}(y_1, \ldots, y_n)\big) \times \\ \times |J(y_1, \ldots, y_n)|. \tag{5.27}$$

Розглянемо випадок, коли вектор $\mathbf{Y} = G(\mathbf{X})$ має вимірність, меншу, ніж вимірність вектора $\mathbf{X}$ $(m < n)$.

Якщо система рівнянь

$$y_k = g_k(x_1, \ldots, x_n), \quad 1 \leqslant k \leqslant m$$

не має єдиного розв'язку відносно координат $(x_1, \ldots, x_m)$ вектора $\mathbf{X}$ ні за якого вибору цих координат, поступимо так. Доповнимо вектор $\mathbf{Y}$ до n-вимірного, ввівши координати $y_{m+1} = x_{m+1}, \ldots, y_n = x_n$. В цьому випадку, як показано вище (формула (5.27))

$$f_{\mathbf{Y}}(y_1, \ldots, y_n) = \sum_{I_k(Y)\in\Delta(Y)} \rho_{\mathbf{X}}\big(g_1^{-1}(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n), \ldots,\\ \ldots, g_n^{-1}(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n\big)\times\\ \times |J(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n)|. \tag{5.28}$$

Тут $g_i^{-1}(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n)$ — розв'язок системи рівнянь

$$y_k = g_k(x_1, \ldots, x_n), \quad 1 \leqslant k \leqslant m$$

відносно координат $(x_1, \ldots, x_m)$ в області можливих значень координат $\mathbf{x} = (x_1, \ldots, x_n)$. Інтегруючи (5.28) за змінними $(x_{m+1}, \ldots, x_m)$, отримуємо щільність вектора $\mathbf{Y}$:

$$\overline{f_{\mathbf{Y}}} = \int\limits_{-\infty}^{+\infty} \mathrm{d}x_{m+1} \ldots \int\limits_{-\infty}^{+\infty} \mathrm{d}x_n\, f_{\mathbf{Y}}(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n). \tag{5.29}$$

Приклад 5.6. Нехай $\mathbf{X} = (\xi_1, \xi_2)$ — випадковий вектор, сукупна щільність розподілу якого є $\rho_{\mathbf{X}}(x_1, x_2)$.

Знайти щільність розподілу випадкової величини:

а) $Y = \xi_1 + \xi_2$;

б) $Y = \dfrac{\xi_1}{\xi_2}$ $(\xi_2 \neq 0)$.

а) Обчислимо щільність розподілу $f_Y(y)$, де $Y = \xi_1 + \xi_2$. Маємо систему рівнянь

$$\begin{cases} y_1 = g_1(x_1, x_2) = x_1 + x_2, \\ y_2 = x_2, \end{cases}$$

з якої слідує, що $\forall (x_1, x_2) \in \mathbb{R}$

$$\begin{cases} x_1 = g_1^{-1}(y_1, y_2) = y_1 - y_2, \\ x_2 = y_2. \end{cases}$$

Якобіан переходу

$$J(y_1,x_2)=\begin{vmatrix} \dfrac{\partial(y_1-y_2)}{\partial y_1} & \dfrac{\partial(y_1-y_2)}{\partial y_2} \\ \dfrac{\partial y_2}{\partial y_1} & \dfrac{\partial y_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

Використовуючи (5.28) та (5.29), маємо

$$f_Y(y)=\int\limits_{-\infty}^{+\infty}\rho_{\mathbf{x}}(y-x_2,x_2)\,\mathrm{d}x_2.$$

Якщо ξ_1 та ξ_2 є незалежними випадковими величинами, тобто $\rho_{\mathbf{x}}(x_1,x_2)=\rho^1_{\xi_1}(x_1)\rho^2_{\xi_2}(x_2)$, то щільність розподілу випадкової величини $Y=\xi_1+\xi_2$ є

$$\begin{aligned} f_Y(y)=\int\limits_{-\infty}^{+\infty}\rho^1_{\xi_1}(y-x_2)\rho^2_{\xi_2}(x_2)\,\mathrm{d}x_2 = \\ =(\rho^1_{\xi_1}*\rho^2_{\xi_2}), \end{aligned}$$

де $(\cdot*\cdot)$ — згортка функцій.

б) В цьому випадку $y_1=\dfrac{x_1}{x_2}$, а $y_2=x_2$. Тому $\forall(x_1,x_2)\in\mathbb{R}$ система рівнянь

$$\begin{cases} y_1=\dfrac{x_1}{x_2}, \\ y_2=x_2 \end{cases}$$

має єдиний розв'язок

$$\begin{aligned} x_1&=y_1\cdot y_2,\\ x_2&=y_2. \end{aligned}$$

Обчислимо якобіан переходу

$$J(y_1,x_2)=\begin{vmatrix} \dfrac{\partial(y_1\cdot y_2)}{\partial y_1} & \dfrac{\partial(y_1\cdot y_2)}{\partial y_2} \\ \dfrac{\partial y_2}{\partial y_1} & \dfrac{\partial y_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2 = x_2.$$

Тому

$$f_Y(y) = \int\limits_{-\infty}^{+\infty} \rho_{\mathbf{x}}(y \cdot x_2, x_2)\, x_2 \,\mathrm{d}x_2.$$

Слід зауважити, що задачу розрахунку щільності розподілу випадкового вектора $\mathbf{Y} = G(\mathbf{X})$ через щільність розподілу випадкового вектора $\mathbf{X}$ у випадку, коли

$$\mathbf{Y} = (\eta_1, \ldots, \eta_m), \quad \mathbf{X} = (\xi_1, \ldots, \xi_n),$$
$$\eta_k = g_k(\xi_1, \ldots, \xi_n), \quad 1 \leqslant k \leqslant m$$

для всіх можливих співвідношень між розмірностями цих векторів ($m = n$, $m < n$, $m > n$) легко розв'язується з використанням узагальнених функцій. Дійсно, нехай $\mathbf{Y} = G(\mathbf{X})$, $\mathbf{Y} = (\eta_1, \ldots, \eta_m)$, $\mathbf{X} = (\xi_1, \ldots, \xi_n)$ і $\eta_k = y_k = g_k(x_1, \ldots, x_n)$, $1 \leqslant k \leqslant m$, де функції $g_k(x_1, \ldots, x_n)$ задовільняють всі умови, що викладені раніше. Величина $\mathbf{Y} = (y_1, \ldots, y_n)$ за довільних значень $(x_1, \ldots, x_n) \in \mathbb{R}$ має єдине можливе значення

$$\mathbf{Y} = \big(g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)\big).$$

Тому умовна щільність розподілу

$$f_2(y_1, \ldots, y_m | x_1, \ldots, x_n) = \prod_{k=1}^{m} \delta\big(y_k - g_k(x_1, \ldots, x_n)\big).$$

Сукупна щільність розподілу випадкових векторів $\mathbf{X} = (\xi_1, \ldots, \xi_n)$ та $\mathbf{Y} = (\eta_1, \ldots, \eta_m)$ задається виразом

$$f(\mathbf{X}, \mathbf{Y}) = \rho_{\mathbf{X}}(x_1, \ldots, x_n) f_2(y_1, \ldots, y_m | x_1, \ldots, x_n).$$

Тому

$$f_{\mathbf{Y}}(x_1, \ldots, y_n) = \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} f(y_1, \ldots, y_m, x_1, \ldots, x_n) \,\mathrm{d}x_1 \ldots \,\mathrm{d}x_n =$$
$$= \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} \rho_{\mathbf{X}}(x_1, \ldots, x_n) \prod_{k=1}^{m} \delta\big(y_k - g_k(x_1, \ldots, x_n)\big) \,\mathrm{d}x_1 \ldots \,\mathrm{d}x_n.$$

ЛЕКЦІЯ 6

Числові характеристики випадкових величин

Для опису випадкових величин ξ, як було викладено в попередніх розділах, найбільш повну інформацію дає функція розподілу

$$F_\xi = \mathbb{P}\{\omega \in \Omega \colon \xi(\omega) < x\}$$

(або закон розподілу, записаний в будь-якій іншій формі, наприклад, щільність розподілу для абсолютно неперервних випадкових величин), для якої важко побудувати ті, чи інші моделі. Однак в багатьох практичних задачах по суті варто знати кілька чисел, які дозволяють описати випадкові величини (однак не так повно, як задання функції розподілу). Ці числа будемо надалі будемо називати *числовими характеристиками* випадкових величин.

Методі побудови числових характеристик випадкових величин ілюструє наступний приклад.

Приклад 6.1. Нехай a — деяка величина, яку необхідно визначити шляхом вимірюванні. За достатньо великої кількості експериментів отримуємо ряд $x_1, \ldots, x_n$ значень величини a. Припустимо, що результат вимірювання x_1 з'явився k_1 разів, x_2 — k_2 разів і т.д. Приймемо за числове a величину

$$\overline{x} = \frac{1}{n} \sum_{k=1}^{n} m_k x_k.$$

При необмеженому збільшенні числа n експериментальних вимірювань величини a відносні частоти $\frac{m_k}{n}$ появи результату експерименту x_k будуть стабілізуватись навколо чисел p_k, які можна трактувати як ймовірності появи в процесі вимірювання величини a значення x_k. Якщо означити значення $\overline{x} = \sum\limits_{k=1}^{N} x_k p_k$ — середнє значення випадкової величини $a(\omega)$ з заданим законом розподілу

$$\mathbb{P}\{\omega \in \Omega\colon a(\omega) = x_k\} = p_k, \quad k = 1 \div N$$

як абстрактну числову характеристику цього закону розподілу, то структура та необхідність дослідження цих величин є зрозумілою. Найважливішою числовою характеристикою скалярної випадкової величини є її *математичне сподівання* (або середнє значення).

§1. Математичне сподівання скалярних випадкових величин: означення та властивості

Нехай $\langle \Omega, \mathfrak{A}, \mathbb{P} \rangle$ — ймовірнісний простір, а $\xi(\omega)$ задана на цьому просторі скалярна випадкова величина. Справедливе наступне означення.

Означення 1. Математичним сподіванням випадкової величини $\xi(\omega)$ називають число (функціонал)

$$\mathrm{M}[\xi] \stackrel{\mathrm{def}}{=} \int\limits_{\Omega} \xi(\omega) \mathbb{P}(\mathrm{d}\omega). \tag{6.1}$$

В означенні (6.1) інтеграл є інтегралом Лебега за ймовірнісною мірою $\mathbb{P}$. В теорії інтегралу Лебега показано, що $\mathrm{M}[\xi]$ існує тоді і тільки тоді, коли існує

$$\mathrm{M}[|\xi|] \stackrel{\mathrm{def}}{=} \int\limits_{\Omega} |\xi(\omega)| \mathbb{P}(\mathrm{d}\omega), \tag{6.2}$$

де
$$|\xi(\omega)| = \xi^+(\omega) + \xi^-(\omega),$$
а
$$\xi^+(\omega) = \max\{0; \xi(\omega)\},$$
$$\xi^-(\omega) = \min\{0; -\xi(\omega)\}$$
— вимірні відносно ймовірнісної міри $\mathbb{P}$ функції.

Якщо для випадкової величини $\xi(\omega)$ задано функцію розподілу
$$F_\xi(x) = \mathbb{P}\{x \in \mathbb{R}, \omega \in \Omega \colon \xi(\omega) < x\},$$
то інтеграл Лебега у виразах (6.1) та (6.2) переходить в інтеграл Лебега-Стільтьєса (інтеграл по мірі, яка асоційована з функцією розподілу $F_\xi(x)$) і справедлива рівність
$$\int_\Omega \xi(\omega)\mathbb{P}(\mathrm{d}\omega) = \int_\mathbb{R} x\,\mathrm{d}F_\xi(x) \tag{6.3}$$
(якщо ці інтеграли існують).

Більше того, якщо $\eta = g(\omega)$, $g(x)$ — вимірна функція, то
$$\int_\Omega g(\xi(\omega))\mathbb{P}(\mathrm{d}\omega) = \int_\mathbb{R} g(x)\,\mathrm{d}F_\xi(x), \tag{6.4}$$
якщо існує
$$\int_\Omega |g(\xi(\omega))|\mathbb{P}(\mathrm{d}\omega).$$
З відомих властивостей інтеграла Лебега-Стільтьєса випливає, що якщо існує $\int\limits_\mathbb{R} g(x)\,\mathrm{d}F_\xi$, то
$$\int_\Omega g(x)\,\mathrm{d}F_\xi(x) = \begin{cases} \int\limits_{-\infty}^{+\infty} g(x)\rho_\xi(x)\,\mathrm{d}x, & \text{якщо } F_\xi(x) \text{ — диференційовна функція} \\ & \text{(існує неперервна щільність } \rho_\xi(x)), \\ \sum\limits_k g(x_k)\big[F_\xi(x_k+0) - F_\xi(x_k-0)\big], & \\ & \text{якщо } F_\xi(x) \text{ — функція стрибків.} \end{cases} \tag{6.5}$$

Формула (6.5) дозволяє записати найбільш вживані в практичному застосуванні формули для розрахунку $\mathrm{M}[\xi]$, а саме:

$$\mathrm{M}[\xi] = \sum_k x_k P\{\xi = x_k\}, \tag{6.6}$$

якщо ξ — дискретна випадкова величина, $P\{\xi = x_k\}$ — її закон розподілу:

$$\mathrm{M}[\xi] = \int\limits_{-\infty}^{+\infty} x\, \rho_\xi(x)\, \mathrm{d}x, \tag{6.7}$$

якщо ξ — абсолютно неперервна випадкова величина, $\rho_\xi(x)$ — її щільність розподілу.

Нехай $\eta = g(\xi)$ — функція випадкової величини, а $g(x)$ — вимірна (борелівська) функція. Тоді (6.5) дозволяє розрахувати

$$\mathrm{M}[g(\xi)] = \sum_j g(x_j) P\{\xi = x_j\}, \tag{6.8}$$

якщо ξ — дискретна випадкова величина;

$$\mathrm{M}[g(\xi)] = \int\limits_{-\infty}^{+\infty} g(x)\, \rho_\xi(x)\, \mathrm{d}x, \tag{6.9}$$

якщо ξ — абсолютно неперервна випадкова величина.

Якщо відома щільність розподілу випадкової величини, чи задано закон розподілу дискретної величини, то дуже часто на практиці виявляється, що математичне сподівання зручніше розрахувати не за формулами (6.6)–(6.9), а за допомогою використання наступної теореми.

Теорема 6.1. *Нехай* $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ *— випадковий вектор зі заданою сукупною щільністю розподілу* $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$*. Якщо інтеграл*

$$\int\limits_{\mathbb{R}^n} |g(x_1, \ldots, x_n)| \rho_{\mathbf{x}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

існує, то існує математичне сподівання випадкової величини $\eta = g(\xi_1, \ldots, \xi_n)$ *і*

$$\mathrm{M}[\eta] = \int\limits_{\mathbb{R}^n} g(x_1, \ldots, x_n) \rho_{\mathbf{x}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n. \tag{6.10}$$

Проілюструємо використання формул (6.6) та (6.7) для розрахунку математичного сподівання на деяких прикладах.

Приклад 6.2. Випадкова величина ξ має біноміальний розподіл:

$$\mathbb{P}\{\omega \colon \xi = k\} = P_n(k) = C_n^k p^k q^{n-k}.$$

Обчислити математичне сподівання $\mathrm{M}[\xi]$.

Так як ξ — дискретна випадкова величина, то згідно з (6.6)

$$\begin{aligned}
\mathrm{M}[\xi] &= \sum_{k=0}^{n} k C_n^k p^k q^{n-k} = \\
&= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} = \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p\, p^k q^{n-k} = \\
&= \sum_{k=1}^{n} n\, p \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} q^{(n-1)-(k-1)} = \\
&= \Big| k-1 = m \Big| = n\, p \sum_{m=0}^{n-1} C_{n-1}^m p^m q^{(n-1)-m} = \\
&= n\, p (p+q)^{n-1} = n\, p.
\end{aligned}$$

Отже, біноміально розподілена випадкова величина має математичне сподівання

$$\mathrm{M}[\xi] = n\,p.$$

Приклад 6.3. А бсолютно неперервна випадкова величина ξ розподілена нормально з параметрами a та σ. Обчислити математичне сподівання $\mathrm{M}[\xi]$.

Оскільки ξ розподілена нормально, то

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

З (6.7) одержуємо

$$\begin{aligned}
\mathrm{M}[\xi] &= \int\limits_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)\mathrm{d}x = \\
&= \left|\begin{array}{c} \dfrac{x-a}{\sigma} = t \\ x = \sigma t + a \\ \mathrm{d}x = \sigma\,\mathrm{d}t \end{array}\right| = \frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{+\infty}(\sigma t + a)\exp\left(-\frac{t^2}{2}\right)\sigma\,\mathrm{d}t = \\
&= \frac{1}{\sqrt{2\pi}}\left[\sigma\int\limits_{-\infty}^{+\infty} t\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t + \int\limits_{-\infty}^{+\infty} a\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t\right] = a,
\end{aligned}$$

бо

$$\int\limits_{-\infty}^{+\infty} t\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t = 0,$$

а

$$\int\limits_{-\infty}^{+\infty}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t = \sqrt{2\pi}.$$

Отже, для нормального закону розподілу $N(a,\sigma)$

$$\mathrm{M}[\xi] = a.$$

Приклад 6.4. Щільність розподілу випадкової величини ξ має вигляд

$$\rho_\xi(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad x \in \mathbb{R}^1$$

(розподіл Коші). Обчислити математичне сподівання $\mathrm{M}[\xi]$.

З (6.7) одержуємо, що

$$\mathrm{M}[\xi] = \int\limits_{-\infty}^{+\infty} \frac{1}{\pi}\frac{x\,\mathrm{d}x}{1+x^2}.$$

Але невласний інтеграл

$$I = \int\limits_{-\infty}^{+\infty} \frac{1}{\pi}\frac{x\,\mathrm{d}x}{1+x^2}$$

є розбіжним (див. [**?**]). Тому для випадкової величини ξ, що має розподіл Коші, математичне сподівання не існує.

Розглянемо властивості математичного сподівання $\mathrm{M}[\xi]$ скалярної випадкової величини $\xi(\omega)$.

1. Нехай випадкова величина $\xi(\omega)$ є майже напевно сталою, тобто $\forall c \in \mathbb{R}$

$$P\{\omega \in \Omega\colon \xi(\omega) = c\} = 1.$$

Тоді

$$\mathrm{M}[\xi] = c. \tag{6.11}$$

Доведення. Нехай $D = \{\omega \in \Omega\colon \xi(\omega) \neq a\}$. Тоді $\mathbb{P}(D) = 0$ і тому, використовуючи властивість адитивності інтеграла Лебега відносно області інтегрування, маємо:

$$\begin{aligned}
\mathrm{M}[\xi] = \int\limits_{\Omega\setminus D} \xi(\omega)\mathbb{P}(\mathrm{d}\omega) + \int\limits_{D} \xi(\omega)\mathbb{P}(\mathrm{d}\omega) = \\
= a\int\limits_{D} \mathbb{P}(\mathrm{d}\omega) = \\
= c\,\mathbb{P}(\overline{c}) + c\,\mathbb{P}(c) = c \cdot 1 = c.
\end{aligned}$$

■

Якщо $c = 0$, то $\mathrm{M}[0] = 0$. Обернене твердження, взагалі кажучи, невірне, тобто, якщо $\mathrm{M}[\xi] = 0$, то ξ не є майже напевно рівною 0. Це твердження ілюструє наступний приклад.

Приклад 6.5. Нехай $\xi(\omega)$ — дискретна випадкова величина, що набуває два значення $x_1 = 1$, $x_2 = -1$. Відповідні ймовірності появи цих значень

$$\mathbb{P}\{\xi = x_1\} = \mathbb{P}\{\xi = x_2\} = p.$$

Обчислимо

$$\mathrm{M}[\xi] = 1 \cdot p + (-1) \cdot p = 0.$$

Однак $\xi(\omega)$ — відмінна від нуля випадкова величина, оскільки приймає ненулеві значення.

Справедливе наступне твердження.

Нехай $\xi \geqslant 0$ і $\mathrm{M}[\xi] = 0$, тоді $\xi = 0$ (майже напевно, яке випливає з властивостей інтегралу Лебега.

2. Лінійність математичного сподівання.

Нехай $(\xi_n)_{n=1}^{N}$ — послідовність випадкових величин, кожна з яких має скінчене математичне сподівання, тобто $\forall n$, $n = 1 \div N$ $\exists \mathrm{M}[\xi_n]$, $|\mathrm{M}[\xi_n]| < +\infty$. Тоді $\forall C_n \in \mathbb{R}^1$, $n = 1 \div N$ справедлива рівність

$$\mathrm{M}\Big[\sum_{n=1}^{N} C_n \xi_n\Big] = \sum_{n=1}^{N} C_n \mathrm{M}[\xi_n]. \tag{6.12}$$

Доведення. Ця властивість випливає з властивості лінійності інтеграла Лебега:

$$\mathrm{M}\Big[\sum_{n=1}^{N} C_n \xi_n\Big] = \int\limits_{\Omega} \sum_{n=1}^{N} C_n \xi_n(\omega) \mathbb{P}(\mathrm{d}\omega) = \\ = \sum_{n=1}^{N} C_n \int\limits_{\Omega} \xi_n(\omega) \mathbb{P}(\mathrm{d}\omega) = \sum_{n=1}^{N} C_n \mathrm{M}[\xi_n],$$

бо за умовою $\forall n$, $n = 1 \div N$, $\exists \mathrm{M}[\xi_n]$ і $|\mathrm{M}[\xi_n]| < +\infty$. ■

3. Як наслідок, з властивості лінійності математичного сподівання отримуємо, що

$$\mathrm{M}[\xi_1 - \xi_2] = \mathrm{M}[\xi_1] - \mathrm{M}[\xi_2],$$

за умови, що їх скінченні математичні сподівання існують $\mathrm{M}[\xi_1]$ та $\mathrm{M}[\xi_2]$ (довести самостійно).

➢ ***Зауваження 6.1.*** *Властивість лінійності дозволяє розглядати комплексні випадкові величини $\xi(\omega) = \mathrm{Re}\,\xi(\omega) + \mathrm{i}\,\mathrm{Im}\,\xi(\omega)$ (комплекснозначні числові функції, які визначені на Ω) і розрахувати їх математичне сподівання як комплексне число*

$$\mathrm{M}[\xi] = \mathrm{M}[\mathrm{Re}\,\xi] + \mathrm{i}\,\mathrm{M}[\mathrm{Im}\,\xi].$$

4. Нехай $I_A(\omega) = \begin{cases} 0, \omega \overline{\in} A, \\ 1, \omega \in A, \end{cases}$ — індикатор випадкової події A. Тоді

$$\mathrm{M}[I_A] = \mathbb{P}(A). \tag{6.13}$$

Дійсно, $I_A(\omega)$ можна розглядати як дискретну випадкову величину ξ, що набуває два значення: $x_1 = 0$ та $x_2 = 1$, з відповідними ймовірностями: $\mathbb{P}\{\xi = 0\} = \mathbb{P}(\overline{A})$, $\mathbb{P}\{\xi = 1\} = \mathbb{P}(A)$. тому, згідно з (6.6)

$$\mathrm{M}[I_A(\omega)] = 0 \cdot \mathbb{P}(\overline{A}) + 1 \cdot \mathbb{P}(A) = \mathbb{P}(A).$$

5. Нехай $\xi(\omega)$ — випадкова величина, що має скінчене математичне сподівання $\mathrm{M}[\xi]$. Тоді

$$\big|\mathrm{M}[\xi]\big| \leqslant \mathrm{M}\big[|\xi|\big]. \tag{6.14}$$

Доведення. Доведемо цю властивість (для прикладу) для абсолютно неперервних величин.

Якщо ξ — абсолютно неперервна випадкова величина зі щільністю $\rho_\xi(x)$, то

$$\big|\mathrm{M}[\xi]\big| = \left|\int\limits_{-\infty}^{+\infty} \rho_\xi(x)\, x\, \mathrm{d}x\right| \leqslant$$
$$\leqslant \int\limits_{-\infty}^{+\infty} |x| \rho_\xi(x)\, \mathrm{d}x = \mathrm{M}\big[|\xi|\big].$$

■

6. Нехай $\xi(\omega)$ — майже напевно ($\mathbb{P}\{\xi \geqslant 0\} = 1$) випадкова величина. Тоді

$$\exists \mathrm{M}[\xi] \geqslant 0. \tag{6.15}$$

Доведемо цю властивість для дискретних випадкових величин.

Доведення. Нехай $\xi(\omega)$ — дискретна випадкова величина зі заданим законом розподілу

$$\mathbb{P}\{\xi = x_k\} = p_k \geqslant 0, \quad k = 1 \div n.$$

Умова невід'ємності $\xi(\omega)$ означає, що $(\forall k \geqslant 1)$, $x_k \geqslant 0$. Тому згідно з (6.6)

$$\mathrm{M}[\xi] = \sum_{k=1}^{n} x_k p_k \geqslant 0,$$

бо $(\forall k \geqslant 1)$, $x_k p_k \geqslant 0$. ■

7. Нехай $\xi_1(\omega)$ та $\xi_2(\omega)$ — випадкові величини зі скінченими математичними сподіваннями $\mathrm{M}[\xi_1]$ та $\mathrm{M}[\xi_2]$. Тоді, якщо майже напевно $\xi_1 \geqslant \xi_2$ (тобто $\mathbb{P}\{\xi_1 \geqslant \xi_2\} = 1$), то

$$\mathrm{M}[\xi_1] \geqslant \mathrm{M}[\xi_2]. \tag{6.16}$$

Доведення. Розглянемо випадкову величину $\eta = \xi_1 - \xi_2$. Так як майже напевно $\eta \geqslant 0$. За властивістю 2 $\exists \mathrm{M}[\eta] = \mathrm{M}[\xi_1] - \mathrm{M}[\xi_2]$, а за властивістю 6 $\mathrm{M}[\eta] \geqslant 0$. Тому $\mathrm{M}[\xi_1] \geqslant \mathrm{M}[\xi_2]$[1]. ■

[1]Властивості 5–7 математичного сподівання є властивостями інтегралу Лебега

8. Нерівність Йєнсена.

Нехай $\xi(\omega)$ — випадкова величина зі скінченим математичним сподіванням $\mathrm{M}[\xi]$. Якщо $g(x)$ — випукла вниз борелівська функція, то

$$g(\mathrm{M}[\xi]) \leqslant \mathrm{M}[g(\xi)]. \tag{6.17}$$

Доведення. Якщо $g(x)$ — випукла вниз функція, то для кожного $x_0 \in \mathbb{R}$ знайдеться таке число $\lambda(x_0)$, що

$$(\forall x \in \mathbb{R})\colon\ g(x) \geqslant g(x_0) + \lambda(x_0)(x - x_0).$$

Покладемо $x = \xi$, а $x_0 = \mathrm{M}[\xi]$ і обчислимо $\mathrm{M}[g(\xi)]$. За властивостями $1, 2$ та 7

$$\begin{aligned}\mathrm{M}[g(\xi)] &\geqslant \mathrm{M}\bigl[g(\mathrm{M}[\xi]) + \lambda(\mathrm{M}[\xi])(\xi - \mathrm{M}[\xi])\bigr] = \\ &= \mathrm{M}\bigl[g(\mathrm{M}[\xi])\bigr] + \lambda(\mathrm{M}[\xi])\bigl(\mathrm{M}[\xi] - \mathrm{M}[\mathrm{M}[\xi]]\bigr) = \\ &= g(\mathrm{M}[\xi]) + \lambda(\mathrm{M}[\xi])\bigl(\mathrm{M}[\xi] - \mathrm{M}[\xi]\bigr) = \\ &= g(\mathrm{M}[\xi]).\end{aligned}$$

Тому

$$\mathrm{M}[g(\xi)] \geqslant g(\mathrm{M}[\xi]).$$

■

9. Нерівність Ляпунова.

Якщо $0 < \alpha \leqslant \beta$, $(\alpha, \beta) \in \mathbb{R}^1$, а ξ — випадкова величина зі скінченним математичним сподіванням $\mathrm{M}[\xi]$, то

$$\bigl(\mathrm{M}\bigl[|\xi|^{\alpha}\bigr]\bigr)^{1/\alpha} \leqslant \bigl(\mathrm{M}\bigl[|\xi|^{\beta}\bigr]\bigr)^{1/\beta}. \tag{6.18}$$

Доведення. Розглянемо функцію $g(x) = x^{\beta/\alpha}$, $x > 0$. Тоді $g'(x) = \frac{\beta}{\alpha} x^{\beta/\alpha - 1}$, а $g''(x) = \frac{\beta}{\alpha}\left(\frac{\beta}{\alpha} - 1\right) x^{\beta/\alpha - 2} \geqslant 0$. Це означає, що функція $g(x) = x^{\beta/\alpha}$ — випукла вниз. Так як $g(x)$ — борелівська функція, то згідно з нерівністю Йєнсена, записаної для випадкової величини $\eta = |\xi|^{\alpha}$, маємо

$$\mathrm{M}\bigl[|\xi|^{\beta}\bigr] \geqslant g\bigl(\mathrm{M}\bigl[|\xi|^{\alpha}\bigr]\bigr)^{\beta/\alpha}.$$

Піднесемо обидві частини нерівності до степеня $1/\beta$ ($\beta > 0$) і отримуємо, що

$$\left(\mathrm{M}\left[|\xi|^{\beta}\right]\right)^{1/\beta} \geqslant \left(\mathrm{M}\left[|\xi|^{\alpha}\right]\right)^{1/\alpha}.$$

■

10. Нерівність Маркова.

Нехай $\xi(\omega)$ — невід'ємна випадкова величина, яка має скінченне математичне сподівання $\mathrm{M}[\xi]$. Тоді

$$(\forall a > 0)\colon P\{\omega\colon \xi(\omega) \geqslant a\} \leqslant \frac{\mathrm{M}[\xi]}{a^2}. \tag{6.19}$$

Доведення. Нехай A — випадкова подія $\{\omega\colon \xi(\omega) \geqslant a\}$, $a > 0$, a $I_A = \begin{cases} 1, & \xi \geqslant a, \\ 0, & \xi < a \end{cases}$ — індикатор множини A. Так як $\xi(\omega)$ — невід'ємна випадкова величина $\xi \geqslant 0$, то $\mathrm{M}[\xi] \geqslant 0$ і $\mathrm{M}[\xi] \geqslant \mathrm{M}[\xi \cdot I_A]$, оскільки $\xi > \xi \cdot I_a$. Справедливий наступний ланцюг нерівностей

$$\begin{aligned} \mathrm{M}[\xi] \geqslant \mathrm{M}[\xi \cdot I_A] \geqslant a \cdot \mathrm{M}[I_A] = a \cdot P(A) = \\ = a \cdot P\{\omega\colon \xi(\omega) \geqslant a\}. \end{aligned}$$

Тому

$$(\forall \varepsilon > 0)\colon P\{\omega\colon \xi(\omega) \geqslant a\} \leqslant \frac{\mathrm{M}[\xi]}{a}.$$

■

Нерівність Маркова дозволяє оцінити ймовірність появи випадкової події $\{\omega\colon \xi(\omega) \geqslant a\}$, якщо відоме математичне сподівання.

Приклад 6.6. Випадкова величина ξ розподілена зі щільністю $\rho_\xi(x)$. Математичне сподівання $\mathrm{M}[\xi] = 2$. Оцінити ймовірність того, що в процесі випробування випадкова величина набуває значення $\xi \in [3, \infty)$.

Використаємо нерівність Маркова. Оскільки $\mathrm{M}[\xi] = 2$, $\varepsilon = 3$, то шукана ймовірність

$$\mathbb{P}\{\omega\colon \xi(\omega) \in [3, +\infty)\} = \mathbb{P}\{\omega\colon \xi(\omega) \geqslant 3\} \leqslant \frac{2}{3}.$$

11. Нерівність Коші-Буняковського.

Нехай ξ_1 та ξ_2 — довільні випадкові величини, що мають скінченні математичні сподівання $\mathrm{M}[\xi_1^2]$ та $\mathrm{M}[\xi_2^2]$. Тоді

$$\big(\mathrm{M}[\xi_1 \cdot \xi_2]\big)^2 \leqslant \mathrm{M}[\xi_1^2] \cdot \mathrm{M}[\xi_2^2]. \tag{6.20}$$

Доведення. Для довільного $b \in \mathbb{R}^1$ виконується нерівність $(b\xi_1 + \xi_2)^2 \geqslant 0$. Обчислимо $\mathrm{M}[(b\xi_1 + \xi_2)^2]$, яке за властивістю 6 невід'ємне, тобто $\mathrm{M}[(b\xi_1 + \xi_2)^2] \geqslant 0$.

$$\begin{aligned} \mathrm{M}[(b\xi_1 + \xi_2)^2] &= \mathrm{M}[b^2\xi_1^2 + 2b\xi_1\xi_2 + \xi_2^2] = \\ &= b^2\mathrm{M}[\xi_1^2] + 2b\mathrm{M}[\xi_1\xi_2] + \mathrm{M}[\xi_2^2] \geqslant 0 \end{aligned}$$

Умова довільності ξ_1 та ξ_2 та існування $\mathrm{M}[\xi_1^2]$ та $\mathrm{M}[\xi_2^2]$ гарантують існування $\mathrm{M}[\xi_1\xi_2]$. Тому отримана нерівність виконуватиметься $\forall b \in \mathbb{R}^1$, якщо дискримінант

$$(2\mathrm{M}[\xi_1\xi_2])^2 - 4\mathrm{M}[\xi_1^2]\mathrm{M}[\xi_2^2] \leqslant 0,$$

що дає

$$(\mathrm{M}[\xi_1\xi_2])^2 \leqslant \mathrm{M}[\xi_1^2]\mathrm{M}[\xi_2^2]$$

■

Зауважимо, що нерівність Коші-Буняковського дає принципову можливість оцінити математичне сподівання добутку випадкових величини за відомими математичними сподіваннями квадратів цих величин.

Якщо ξ_1 та ξ_2 — довільні *незалежні* випадкові величини, то справедлива наступна властивість математичного сподівання.

12. Нехай ξ_1 та ξ_2 — довільні *незалежні* випадкові величини зі скінченними математичними сподіваннями ($\exists \mathrm{M}[|\xi_1|] < +\infty$, $\exists \mathrm{M}[|\xi_2|] < +\infty$). Тоді

$$\exists \mathrm{M}[\xi_1 \cdot \xi_2] = \mathrm{M}[\xi_1] \cdot \mathrm{M}[\xi_2]. \tag{6.21}$$

Доведення. Використаємо (6.10) для функції $g(\xi_1, \xi_2) = \xi_1 \cdot \xi_2$. Оскільки випадкові величини ξ_1 та ξ_2 — незалежні, то сукупна щільність розподілу

$$\rho_{\xi_1\cdot\xi_2}(x_1,x_2)=\rho^{(1)}_{\xi_1}(x_1)\rho^{(2)}_{\xi_2}(x_2)$$

і тому

$$\begin{aligned}\mathrm{M}[\xi_1\cdot\xi_2] &= \int\limits_{\mathbb{R}} x_1\cdot x_2\cdot\rho_{\xi_1\cdot\xi_2}(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x_1\cdot x_2\cdot\rho^{(1)}_{\xi_1}(x_1)\cdot\rho^{(2)}_{\xi_2}(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \\ &= \int\limits_{-\infty}^{+\infty} x_1\cdot\rho^{(1)}_{\xi_1}(x_1)\,\mathrm{d}x_1\cdot\int\limits_{-\infty}^{+\infty} x_2\cdot\rho^{(2)}_{\xi_2}(x_2)\,\mathrm{d}x_2 = \\ &= \mathrm{M}[\xi_1]\cdot\mathrm{M}[\xi_2],\end{aligned}$$

бо $\mathrm{M}[\xi_1]$ та $\mathrm{M}[\xi_2]$ існують і скінченні. ■

Використовуючи метод математичної індукції (чи формулу (6.10)) це правило множення математичних сподівань узагальнюється на довільну кількість *взаємно незалежних випадкових величин* $\{\xi_i\}_{i=1}^n$ — *взаємно незалежні* випадкові величини зі скінченними математичними сподіваннями $(\forall i,\ i=1\div n)$ $\exists\mathrm{M}[|\xi_i|]<+\infty$, тоді $\exists\mathrm{M}[|\prod\limits_{i=1}^n\xi_i|]<+\infty$ і

$$\mathrm{M}\Big[\prod_{i=1}^n\xi_i\Big]=\mathrm{M}[\xi_1\cdot\xi_2\cdot\ldots\cdot\xi_n]=\prod_{i=1}^n\mathrm{M}[\xi_i]. \tag{6.22}$$

§2. Дискретні випадкові величини

Нехай $\langle\Omega,\mathfrak{A},\mathbb{P}\rangle$ — ймовірнісний простір , а $\xi(\omega)$ — скалярна випадкова величина, яка задана на цьому просторі, для якої існує скінчене математичне сподівання $\mathrm{M}[\xi]$. Природньо розглянути замість випадкової величини ξ її відхилення від $\mathrm{M}[\xi]$

$$\overset{\circ}{\xi}=\xi-\mathrm{M}[\xi]. \tag{6.23}$$

Означення 2. Випадкову величину $\overset{\circ}{\xi}$ називають *центрованою випадковою величиною.*

Зауважимо, що $\mathrm{M}[\overset{\circ}{\xi}] \equiv 0$. Дійсно, завжди

$$\begin{aligned}\mathrm{M}[\overset{\circ}{\xi}] &= \mathrm{M}\big[\xi - \mathrm{M}[\xi]\big] = \mathrm{M}\big[\xi - \mathrm{M}[\mathrm{M}[\xi]\big] = \\ &= \mathrm{M}[\xi] - \mathrm{M}[\xi] \equiv 0.\end{aligned}$$

Цей результат є зрозумілим, він вказує, що випадкова величина "одинаково часто" відхиляється від свого математичного сподівання як вправо, так і вліво. Тому математичне сподівання $\mathrm{M}[\overset{\circ}{\xi}]$ не відображає поведінку $\overset{\circ}{\xi}$ і не може бути віднесене до числових характеристик випадкової величини $\overset{\circ}{\xi}$.

Введемо до розгляду випадкову величину $\overset{\circ}{\xi}{}^2$ (а разом і для ξ) величину

$$\mathrm{M}\big[\overset{\circ}{\xi}{}^2\big] = \mathrm{M}\big[(\xi - \mathrm{M}[\xi])^2\big].$$

Оскільки $(\xi - \mathrm{M}[\xi])^2 \leqslant 2(\xi^2 + \mathrm{M}^2[\xi])$, то якщо існує $\mathrm{M}[\xi^2]$, то існує $\mathrm{M}[(\xi - \mathrm{M}[\xi])^2] = \mathrm{M}[\overset{\circ}{\xi}{}^2]$ і тому зручно розглядати як числову (відмінну від нуля) характеристику випадкової величини $\overset{\circ}{\xi}$ математичне сподівання $\mathrm{M}[\overset{\circ}{\xi}{}^2]$.

Означення 3. *Дисперсією* випадкової величини $\xi(\omega)$ називають математичне сподівання квадрату відхилення цієї величини від її математичного сподівання $\mathrm{M}[\xi]$:

$$\mathrm{D}[\xi] = \mathrm{M}[\overset{\circ}{\xi}{}^2] = \mathrm{M}\big[(\xi - \mathrm{M}[\xi])^2\big]. \tag{6.24}$$

Якщо скористатися властивостями 1–2 математичного сподівання, то вираз (6.24) можна привести до наступного:

$$\begin{aligned}\mathrm{D}[\xi] &= \mathrm{M}\big[(\xi - \mathrm{M}[\xi])^2\big] = \mathrm{M}\big[\xi^2 - 2\xi\,\mathrm{M}[\xi] + \mathrm{M}^2[\xi]\big] = \\ &= \mathrm{M}\big[\xi^2\big] - \mathrm{M}^2[\xi]\big].\end{aligned} \tag{6.25}$$

Якщо випадкова величина $\xi(\omega)$ описується заданою функцією розподілу $F_\xi(x)$, то поклавши в формулі (6.5) $g(\xi) = (\xi - \mathrm{M}[\xi])^2$ отримуємо зручні для практичних застосувань формули для розрахунку $\mathrm{D}[\xi]$, а саме:

$$\mathrm{D}[\xi] = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} \big(x - \mathrm{M}[\xi]\big)^2 \rho_\xi(x)\,\mathrm{d}x, & \text{якщо } F_\xi(x) \text{ є диференційовною функцією,} \\ \displaystyle\sum_k \big(x_k - \mathrm{M}[\xi]\big) p_k, & \text{якщо } F_\xi(x) \text{ — функція стрибків} \\ \qquad p_k = F_\xi(x_k + 0) - F_\xi(x_k - 0). \end{cases} \tag{6.26}$$

Приклад 6.7. Дискретна випадкова величина ξ розподілена за законом Пуассона

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

Оскільки $\mathrm{M}[\xi] = \lambda$ (обчисліть самостійно), то використаємо формулу (6.25):

$$\mathrm{D}[\xi] = \mathrm{M}[\xi^2] - \mathrm{M}^2[\xi]$$

і маємо:

$$\begin{aligned} \mathrm{M}[\xi^2] &= \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} = \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} = \\ &= \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!}\,\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} ((k-1)+1) \frac{\lambda^k}{(k-1)!} = \\ &= \mathrm{e}^{-\lambda} \sum_{k=2}^{\infty} \frac{(k-1)}{(k-1)!} \lambda^k + \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} \lambda^k \frac{1}{(k-1)!} = \\ &= \lambda^2 \mathrm{e}^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda\,\mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \\ &= \lambda^2 \mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda} + \lambda\,\mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda} = \lambda^2 + \lambda, \end{aligned}$$

а тому

$$\mathrm{D}[\xi] = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Для закону *розподілу Пуассона*

$$\mathrm{M}[\xi] = \mathrm{D}[\xi] = \lambda,$$

де λ — параметр розподілу.

Приклад 6.8. Випадкова величина ξ розподілена нормально з параметрами $a = \mathrm{M}[\xi]$ та σ. Обчислити дисперсію.

За формулою (6.26) маємо

$$\begin{aligned}
\mathrm{D}[\xi] &= \int\limits_{-\infty}^{+\infty} \bigl(x - \mathrm{M}[\xi]\bigr)^2 \rho_\xi(x)\,\mathrm{d}x = \\
&= \int\limits_{-\infty}^{+\infty} (x-a)^2 \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)\,\mathrm{d}x = \\
&= \left| \begin{array}{c} \dfrac{x-a}{\sqrt{2}\sigma} = t \\ \mathrm{d}x = \sqrt{2}\sigma\,\mathrm{d}t \end{array} \right| = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} 2\sigma^2 t^2 \exp\left(-t^2\right) \sqrt{2}\sigma\,\mathrm{d}t = \\
&= \frac{\sigma^2}{\sqrt{\pi}} \int\limits_{-\infty}^{+\infty} 2t^2 \exp\left(-t^2\right)\,\mathrm{d}t = \left| \begin{array}{c} \text{інтегруючи частинами} \\ u = t \\ \mathrm{d}v = 2t\exp(-t^2)\,\mathrm{d}t \end{array} \right| = \\
&= -\frac{\sigma^2}{\sqrt{\pi}}\, t\exp(-t^2)\Big|_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{\pi}} \int\limits_{-\infty}^{+\infty} \exp(-t^2)\,\mathrm{d}t = \\
&= \frac{\sigma^2}{\sqrt{\pi}} \sqrt{\pi} = \sigma^2.
\end{aligned}$$

$$\mathrm{D}[\xi] = \sigma^2, \quad \sigma[\xi] = \sigma.$$

Таким чином, з'ясовано теоретико-ймовірнісний зміст параметрів a та σ нормального розподілу $N(a, \sigma)$: a — математичне сподівання, σ^2 — дисперсія. Можна сказати, що нормальний закон розподілу повністю визначається математичним сподіванням і дисперсією.

Розглянемо властивості дисперсії.

1. Дисперсія довільної випадкової величини невід'ємна.

Доведення. Ця властивість безпосередньо випливає з означення та властивості 6 математичного сподівання. Дійсно, $(\xi - \mathrm{M}[\xi])^2 \geqslant 0 \Rightarrow \mathrm{M}[(\xi - \mathrm{M}[\xi])^2] \geqslant 0$. ■

І тому можна ввести до розгляду числову характеристику

$$\sigma[\xi] = \sqrt{\mathrm{D}[\xi]}, \tag{6.27}$$

яку називають *середнім квадратичним відхиленням (стандартом)* — мірою розсіяння значень випадкової величини навколо математичного сподівання $\mathrm{M}[\xi]$.

2. Нехай випадкова величина $\xi(\omega)$ є майже напевно стала, тобто $(\forall c \in \mathbb{R})\colon \mathbb{P}\{\xi = c\} = 1$. Тоді

$$P[c] = 0.$$

Доведення. Скориставшись означенням (6.23) та властивістю 1 математичного сподівання, отримуємо, що

$$\mathrm{D}[c] = \mathrm{M}\big[(c - \mathrm{M}[c])^2\big] = \mathrm{M}\big[(c - c)^2\big] = \\ = \mathrm{M}[0] = 0.$$

■

Має місце і обернене твердження:
якщо $\mathrm{D}[\xi] = 0$, то $\forall c \in \mathbb{R}\colon \mathbb{P}\{\omega \in \Omega\colon \xi() = c\} = 1$.

Доведення. З означення (6.24) дисперсії $\mathrm{D}[\xi]$ випливає, що

$$\mathrm{D}[\xi] = \mathrm{M}\big[(\xi - \mathrm{M}[\xi])^2\big] = 0.$$

Так як $(\xi - \mathrm{M}[\xi])^2 \geqslant 0$, то

$$\mathrm{M}\big[(\xi - \mathrm{M}[\xi])^2\big] = \int\limits_{\Omega} (\xi - \mathrm{M}[\xi])^2 \mathbb{P}(\mathrm{d}\omega) = 0 \quad \Longrightarrow \\ \Longrightarrow \quad \mathbb{P}\{\omega \in \Omega\colon \xi(\omega) = c\} = 1$$

(властивість інтеграла Лебега). ■

3. Нехай c — довільна дійсна стала, а ξ — випадкова величина зі заданою дисперсією $\mathbb{D}[\xi]$. Тоді

$$(\forall c \in \mathbb{R}^1) \colon \exists \mathrm{D}[c\,\xi] = c^2 \mathrm{D}[\xi]. \tag{6.28}$$

Доведення. Використовуючи (6.24) та властивість лінійності математичного сподівання (властивість 2), отримуємо:

$$\begin{aligned} \mathrm{D}[c\,\xi] &= \mathrm{M}\big[(c\,\xi - \mathrm{M}[c\,\xi])^2\big] = \mathrm{M}\big[(c\,\xi - c\,\mathrm{M}[\xi])^2\big] = \\ &= \mathrm{M}\big[c^2(\xi - \mathrm{M}[\xi])^2\big] = c^2\mathrm{M}\big[(\xi - \mathrm{M}[\xi])^2\big] = \\ &= c^2\mathrm{D}[\xi]. \end{aligned}$$

Це доводить (6.28) і забезпечує існування $\mathrm{D}[c\,\xi]$. ■

4. Нехай ξ_1 та ξ_2 — незалежні випадкові величини зі скінченними дисперсіями $\mathrm{D}[\xi_1]$ та $\mathrm{D}[\xi_2]$. Тоді

$$\mathrm{D}[\xi_1 + \xi_2] = \mathrm{D}[\xi_1] + \mathrm{D}[\xi_2]. \tag{6.29}$$

Доведення. Так як ξ_1 та ξ_2 — незалежні випадкові величини, то (див. (6.21))

$$\mathrm{M}[\xi_1 \cdot \xi_2] = \mathrm{M}[\xi_1] \cdot \mathrm{M}[\xi_2]$$

і тому

$$\begin{aligned} \mathrm{D}[\xi_1 + \xi_2] &= \mathrm{M}\big[(\xi_1 + \xi_2)^2\big] - \big(\mathrm{M}[\xi_1 + \xi_2]\big)^2 = \\ &= \mathrm{M}\big[\xi_1^2 + 2\xi_1\xi_2 + \xi_2^2\big] - \big(\mathrm{M}[\xi_1] + \mathrm{M}[\xi_2]\big)^2 = \\ &= \mathrm{M}\big[\xi_1^2\big] + 2\mathrm{M}[\xi_1\xi_2] + \mathrm{M}\big[\xi_2^2\big] - \\ &\quad - \big(\mathrm{M}[\xi_1]\big)^2 - 2\mathrm{M}[\xi_1]\mathrm{M}[\xi_2] - \big(\mathrm{M}[\xi_2]\big)^2 = \\ &= \mathrm{M}\big[\xi_1^2\big] - \mathrm{M}^2[\xi_1] + \mathrm{M}\big[\xi_2^2\big] - \mathrm{M}^2[\xi_2] = \\ &= \mathrm{D}[\xi_1] + \mathrm{D}[\xi_2]. \end{aligned}$$

■

5. Якщо $\{\xi_n\}_{n=1}^{N}$ — послідовність попарно незалежних випадкових величини зі скінченними дисперсіями $\forall n$, $\exists \mathrm{D}[\xi_n]$, то

$$\mathrm{D}\Big[\sum_{n=1}^{N}\xi_n\Big] = \sum_{n=1}^{N}\mathrm{D}[\xi_n]. \tag{6.30}$$

Доведення. Доведення є прямим узагальненням властивості 4 дисперсії, оскільки для попарно незалежних випадкових величин

$$(\forall i \neq k)\colon \mathrm{M}[\xi_i \cdot \xi_k] = \mathrm{M}[\xi_i] \cdot \mathrm{M}[\xi_k]$$

і тому

$$\begin{aligned}
\mathrm{D}\Big[\sum_{n=1}^{N}\xi_n\Big] &= \mathrm{M}\Big[\Big(\sum_{n=1}^{N}\xi_n\Big)^2\Big] - \mathrm{M}^2\Big[\sum_{n=1}^{N}\xi_n\Big] = \\
&= \mathrm{M}\Big[\sum_{n=1}^{N}\xi_n^2 + 2\sum_{i<k}\xi_i\xi_k\Big] - \\
&\quad - \sum_{n=1}^{N}\mathrm{M}^2[\xi_n] - 2\sum_{i\leqslant k}\mathrm{M}[\xi_i]\mathrm{M}[\xi_k] = \\
&= \sum_{n=1}^{N}\big(\mathrm{M}\big[\xi_n^2\big] - \mathrm{M}^2[\xi_n]\big) = \sum_{n=1}^{N}\mathrm{D}[\xi_n].
\end{aligned}$$

■

Приклад 6.9. Обчислити $\mathrm{D}[\xi]$, де ξ — біноміально розподілена випадкова величина.

Нехай p — параметр біноміального розподілу і μ_n — кількість успіхів в схемі Бернуллі (n — кількість випробувань). Якщо I_A — індикатор множини A, то

$$\xi = \mu_n = \xi_1 + \ldots + \xi_n,$$

де $\xi_i = I_{A_1^{(i)}}$, $i = 1, 2, \ldots, n$, $A_1^{(i)}$ — випробування з номером i завершилось успіхом. Оскільки події ξ_i попарно незалежні, то

$$\mathrm{D}\mu_n = \sum_{i=1}^{n}\mathrm{D}[\xi_i],$$

а

$$\mathrm{D}[\xi_i] = \mathrm{D}[I_{A_1^{(i)}}] = \mathrm{M}[I^2_{A_1^{(i)}}] - \mathrm{M}^2[I_{A_1^{(i)}}] = p - p^2 = p(1-p).$$

Тому

$$\mathrm{D}[\xi] = \sum_{i=1}^{n} p(1-p) = np(1-p) = npq.$$

Тут $q = 1 - p$.

Зауваження.

1. У формулювання теореми Лапласа-Муавра входить випадкова величина

$$\frac{\mu_n - np}{\sqrt{npq}} = \frac{\mu_n - \mathrm{M}[\mu_n]}{\sqrt{\mathrm{D}[\mu_n]}}.$$

Таке лінійне перетворення випадкових величин дуже часто використовують з того, що якщо

$$\eta = \frac{\xi - \mathrm{M}[\xi]}{\sqrt{\mathrm{D}[\xi]}},$$

то $\mathrm{M}[\eta] = 0$, $\mathrm{D}[\eta] = 1$.

2. Якщо випадкові величини $\xi_1, \ldots, \xi_n$ *залежні*, то

$$\mathrm{D}\Big[\sum_{i=1}^{n} \xi_i\Big] = \sum_{i=1}^{n} \mathrm{D}[\xi_i] + \sum_{i \neq j} \mathrm{M}[\xi_i \xi_j]. \tag{6.31}$$

§3. Моменти вищих порядків. Деякі інші числові характеристики випадкових величин

В багатьох практичних задачах моментів першого та другого порядків (математичного сподівання та дисперсії) виявляється недостатньо. Тому для більш повного опису випадкових величин використовують моменти вищих порядків. Нехай $a \in \mathbb{R}$ — дійсне число, а $k \geqslant 0$ — ціле.

Означення 4. Моментом k-го порядку випадкової величини ξ відносно точки називають число (якщо воно існує)

$$\nu_k(a) = \mathrm{M}\big[(\xi - a)^k\big]. \tag{6.32}$$

Якщо $a = 0$, то момент називають *початковим моментом* ν_k. Зрозуміло, що $\nu_1 = \mathrm{M}[\xi]$.

Означення 5. Центральним моментом μ_k k-го порядку випадкової величини ξ називають момент k-го порядку відносно точки $a = \mathrm{M}[\xi]$ (або початковий момент для k-ої степені центрованої випадкової величини)

$$\mu_k = \nu_k\big(\mathrm{M}[\xi]\big) = \mathrm{M}\big[\overset{\circ}{\xi}{}^k\big] = \mathrm{M}\big[(\xi - \nu_1)^k\big]. \tag{6.33}$$

Розкриваючи вираз у правій частині (6.33), отримуємо зв'язки між початковими та центральними моментами, наприклад,

$$\begin{aligned}
&\mu_1 = 0,\\
&\mu_2 = \nu_2 - \nu_1^2,\\
&\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3,\\
&\ldots \quad \ldots\ldots\ldots\ldots\ldots
\end{aligned} \tag{6.34}$$

Якщо випадкова величина ξ задається своєю функцією розподілу $F_\xi(x)$, то моменти ν_k та μ_k розраховуються за формулами (6.8) чи (6.9), а саме:

$$\nu_k = \begin{cases} \displaystyle\sum_i x_i^k\, \mathbb{P}\{\xi = x_i\}, \\ \displaystyle\int_{-\infty}^{+\infty} x^k \rho_\xi(x)\,\mathrm{d}x, \end{cases} \tag{6.35}$$

а

$$\mu_k = \begin{cases} \displaystyle\sum_i (x_i - \nu_1)^k\, \mathbb{P}\{\xi = x_i\}, \\ \displaystyle\int_{-\infty}^{+\infty} (x - \nu_1)^k \rho_\xi(x)\,\mathrm{d}x. \end{cases} \tag{6.36}$$

Окрім моментів k-го порядку використовують ще інші числові характеристики.

Медіаною розподілу $F_\xi(x)$ називають таке значення аргументу $x = m$, для якого виконується нерівність

$$F_\xi(m) \leqslant \frac{1}{2} \leqslant F_\xi(m+0). \tag{6.37}$$

Кожний розподіл $F_\xi(x)$ має принаймні *одну медіану*, бо функція $F_\xi(x)$ зростає від 0 до 1. Якщо криві $y = F_\xi(x)$ та $y = 1/2$ мають спільний відрізок, то абсциса кожної точки цього відрізка може вважатись медіаною.

Наприклад, для нормального розподілу $N(a, \sigma)$ $m = a$.

Аналогічно поняттю медіани m для довільного $p \in (0, 1)$ означають *квантиль* розподілу порядку p.

Якщо $F_\xi(x)$ неперервна, то квантиль порядку p є коренем рівняння $F_\xi(x) = p$. Очевидно, що медіана — це квантиль порядку $1/2$. Знання квантилів для кількох значень p дає можливість отримати певну уяву про поведінку $F_\xi(x)$. На практиці часто користуються квантилями для $p = 0.1; 0.2; 0.3; \ldots; 0.9$ (їх називають *децилями*) та $p = \frac{1}{4}; \frac{3}{4}$ (їх називають *квартилями*).

Для абсолютно неперервних випадкових величин, які задані щільністю розподілу $\rho_\xi(x)$, вводять поняття *моди розподілу.*

Модою розподілу називають значення $x = x_0$, за яких $\rho_\xi(x)$ досягає максимуму.

Так, наприклад, нормальний розподіл $N(a, \sigma)$ має єдину моду $x = a = \mathrm{M}[\xi]$. Такі розподіли називають *унімодальними.*

Якщо розподіл $F_\xi(x)$ є *симетричним* відносно точки a, тобто виконується рівність

$$F_\xi(x + a) = 1 - F_\xi(a - x + 0)$$

(для абсолютно неперервних випадкових величин це означає, що графік щільності розподілу є симетричним відносно точки a), то математичне сподівання $\mathrm{M}[\xi]$ (якщо воно існує) дорівнює медіані і дорівнює a, тобто

$$\mathrm{M}[\xi] = m = a.$$

для унімодальних розподілів

$$\mathrm{M}[\xi] = m = Q = a.$$

Симетричні розподіли мають особливість:

$$\forall k > 0, \quad \mu_{2k+1} = \mathrm{M}\big[(\xi - \mathrm{M}[\xi])^{2k+1}\big] \equiv 0.$$

Тому кожний відмінний від нуля центральний момент непарного порядку характеризує ступінь асиметрії.

За характеристику несиметричності розподілу приймають *асиметрію*

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\mathrm{D}^{3/2}} = \frac{\mu_3}{\mu_2^{3/2}}. \tag{6.38}$$

Центральний момент четвертого порядку μ_4 за заданої дисперсії служить характеристикою питомої ваги великих відхилень від математичного сподівання, що в свою чергу визначає характер максимуму в точці $\mathrm{M}[\xi]$ симетричного розподілу ("гостровершинність" чи "плосковершинність" кривої розподілу $F_\xi(x)$). Тому поведінку кривої розподілу в околі точки $\mathrm{M}[\xi]$ описують *ексцесом* розподілу випадкової величини ξ

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3. \tag{6.39}$$

Така структура ексцесу пов'язана з тим, що для нормального закону розподілу $\mu_4 = 3\mu_2^2$ (переконайтесь самостійно), а тому ексцес нормального розподілу дорівнює нулю.

§4. Числові характеристики випадкових векторів. Коваріація і коефіцієнт кореляції

Нехай $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ — n-вимірний випадковий вектор, "координати" якого ξ_i, $i = 1 \div n$ — випадкові величини, задані на ймовірнісному просторі $\langle \Omega, \mathfrak{A}, \mathbb{P} \rangle$. Будемо розглядати випадок, коли відома сукупна функція розподілу $F_{\mathbf{x}}(x_1, \ldots, x_n)$ вектора $\mathbf{x}$. Справедливе наступне означення.

Означення 6. Математичним сподіванням випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ називають вектор

$$\mathrm{M}[\mathbf{x}] = (\mathrm{M}[\xi_1], \ldots, \mathrm{M}[\xi_n]),$$

якщо існують скінченні математичні сподівання $\mathrm{M}[\xi_i]$ компонент ξ_i, тобто $(\forall i = 1 \div n)\ |\mathrm{M}[\xi_i]| < +\infty$.

Якщо задано сукупну функцію розподілу $\mathbf{x} = (\xi_1, \ldots, \xi_n)$, то тим задано або сукупний закон розподілу $\mathbb{P}\{\omega \in \Xi \colon \xi_i = x_i, i = 1 \div n\}$ (якщо вектор $\mathbf{x}$ — дискретний випадковий вектор), або сукупна щільність розподілу $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$ (якщо компоненти ξ_i вектора $\mathbf{x}$ абсолютно неперервні випадкові величини). Знаючи ці характеристики, можна розрахувати математичне сподівання вектора $\mathbf{x}$. Проілюструємо це для випадку $n = 2$.

За означенням математичного сподівання скалярної випадкової величини (наприклад, компоненти ξ_1)

$$\mathrm{M}[\xi_1] = \begin{cases} \int\limits_{-\infty}^{\infty} x\, \rho_{\xi_1}(x)\, \mathrm{d}x, \\ \sum\limits_{x_i} \mathbb{P}(\xi_1 = x_i), \end{cases}$$

де $\rho_{\xi_1}(x)$ (чи $\mathbb{P}(\xi_1 = x_i)$) — маргінальна щільність (закон розподілу) компоненти ξ_1. За формулами (**??**) маємо:

$$\rho_{\xi_1}(x_1) = \int\limits_{-\infty}^{\infty} \rho_{\mathbf{x}}(x_1, x_2)\, \mathrm{d}x_2,$$
$$\mathbb{P}(\xi_1 = x_i) = \sum_{x_2^{(j)}} \mathbb{P}\left\{\xi_1 = x_i; \xi_2 = x_2^{(j)}\right\},$$

а тому

$$\mathrm{M}[\xi_1] = \begin{cases} \displaystyle\int\limits_{-\infty}^{\infty} x_1\, \rho_{\mathbf{x}}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2, \\ \displaystyle\sum_{x_1^{(i)}, x_2^{(j)}} x_1^{(i)} \mathbb{P}\{\xi_1 = x_1^{(i)}, \xi_2 = x_2^{(j)}\}, \end{cases} \tag{6.40}$$

$$\mathrm{M}[\xi_2] = \begin{cases} \displaystyle\iint\limits_{-\infty}^{\infty} x_2\, \rho_{\mathbf{x}}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2, \\ \displaystyle\sum_{x_1^{(i)}, x_2^{(j)}} x_2^{(j)} \mathbb{P}\{\xi_1 = x_1^{(i)}, \xi_2 = x_2^{(j)}\}. \end{cases} \tag{6.41}$$

Формули (6.40) та (6.41) легко узагальнюються на випадок $n > 2$, а саме:
$\mathbf{x} = (\xi_1, \ldots, \xi_n)$ — дискретний вектор, то

$$\mathrm{M}[\xi_i] = \sum_{j_1, \ldots, j_n} x_{j_i}\, \mathbb{P}\{\omega \in \Omega\colon \xi_i = x_j, 1 \leqslant j \leqslant n\}, \quad i = 1 \div n \tag{6.42}$$

$\mathbf{x} = (\xi_1, \ldots, \xi_n)$ — абсолютно неперервний вектор, то

$$\mathrm{M}[\xi_i] = \int\limits_{\mathbb{R}^n} x_i\, \rho_{\mathbf{x}}(x_1, \ldots, x_n) \prod_{j \neq i} \mathrm{d}x_j, \quad i = 1 \div n. \tag{6.43}$$

Приклад 6.10. Знайти математичне сподівання вектора $\mathbf{x} = (\xi_1, \xi_2)$, сукупна щільність розподілу якого є

$$\rho_{\mathbf{x}}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{(x_1-a)^2}{2\sigma_1^2} + \right.\right.$$
$$\left.\left. + \frac{(x_2-b)^2}{2\sigma_2^2} - 2\rho \frac{(x_1-a)(x_2-b)}{\sigma_1\sigma_2} \right) \right\},$$
$$a, b \in \mathbb{R}, \quad \sigma_1, \sigma_2 > 0, \quad |\rho| < 1.$$

Розрахуємо маргінальну щільність $\rho_{\xi_1}(x_1)$ компоненти ξ_1. Маємо

$$\rho_{\xi_1}(x_1) = \int\limits_{-\infty}^{\infty} \rho_{\mathbf{x}}(x_1, x_2)\,\mathrm{d}x_2 =$$

$$= \left| \begin{array}{l} \text{виділимо у показнику експоненти} \\ \text{повний квадрат і зробимо заміну} \\ t = \dfrac{1}{\sqrt{2(1-p^2)}}\left(\dfrac{x_2-b}{\sigma_2} - p\dfrac{x_1-a}{\sigma_1}\right) \end{array} \right| =$$

$$= \frac{\exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2}\right\}}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}} \int\limits_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-p^2)}\left(\frac{x_2-b}{\sigma_2} - p\frac{x_1-a}{\sigma_1}\right)^2\right\} \mathrm{d}x_2 =$$

$$= \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_1}\exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2}\right\} \int\limits_{-\infty}^{\infty} \exp(-t^2)\,\mathrm{d}t.$$

Отже,

$$\rho_{\xi_1}(x_1) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_1}\exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2}\right\} \int\limits_{-\infty}^{\infty} \exp(-t^2)\,\mathrm{d}t.$$

Аналогічно отримуємо

$$\rho_{\xi_2}(x_2) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_2}\exp\left\{-\frac{(x_2-b)^2}{2\sigma_2^2}\right\} \int\limits_{-\infty}^{\infty} \exp(-t^2)\,\mathrm{d}t.$$

Тобто компоненти ξ_1 та ξ_2 розподілені нормально. Це означає, що

$$\mathrm{M}[\xi_1] = a, \quad \mathrm{M}[\xi_2] = b,$$

а математичне сподівання вектора $\mathbf{x}$ є числовий вектор (a, b).

Означення 7. Дисперсією (дисперсійною матрицею) випадкового вектора $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ називають матрицю $\mathbb{K} = \|k_{ij}\|$ ($\tau(\mathbb{K}) = n \times n$), де

$$k_{ij} = \operatorname{cov}(\xi_i, \xi_j) = \mathrm{M}\big[(\xi_i - \mathrm{M}[\xi_i])(\xi_j - \mathrm{M}[\xi_j])\big], \tag{6.44}$$

якщо існує скінченне математичне сподівання компонент ξ ($\forall i$, $\exists \mathrm{M}[|\xi_i|] < +\infty$)

Дисперсійна матриця $\mathbb{K}$ має наступні властивості, які випливають з означення (6.44).

1) Матриця $\mathbb{K}$ є симетричною матрицею: $k_{ij} = k_{ji}$.
2) $\mathrm{cov}(c\,\xi_i, \xi_j) = c\,\mathrm{cov}(\xi_i, \xi_j)$, $c \in \mathbb{R}$. Дійсно, як випливає з (6.44)

$$\begin{aligned}\mathrm{cov}(c\,\xi_i, \xi_j) &= \mathrm{M}\big[(c\,\xi_i - \mathrm{M}[c\,\xi_i])(\xi_j - \mathrm{M}[\xi_j])\big] = \\ &= \mathrm{M}\big[c(\xi_i - \mathrm{M}[\xi_i])(\xi_j - \mathrm{M}[\xi_j])\big] = \\ &= c\,\mathrm{M}\big[(\xi_i - \mathrm{M}[\xi_i])(\xi_j - \mathrm{M}[\xi_j])\big] = c\,\mathrm{cov}(\xi_i, \xi_j).\end{aligned}$$

3) $\mathrm{cov}(\xi_i, \xi_i) = \mathrm{D}[\xi_i]$. Обчислимо

$$\mathrm{cov}(\xi_i, \xi_i) = \mathrm{M}\big[(\xi_i - \mathrm{M}[\xi_i])^2\big] = \mathrm{D}[\xi_i].$$

Якщо вимірність вектора $\mathbf{x}$ $n = 2$, то маємо практичну формулу обчислення дисперсії суми двох випадкових величини:

$$\mathrm{D}[\xi_1 + \xi_2] = \mathrm{D}[\xi_1] + \mathrm{D}[\xi_2] + \mathrm{cov}(\xi_1, \xi_2). \qquad (6.45)$$

Дійсно, використовуючи властивість лінійності математичного сподівання, маємо:

$$\begin{aligned}\mathrm{D}[\xi_1 + \xi_2] &= \mathrm{M}\left[((\xi_1 + \xi_2) - \mathrm{M}[\xi_1 + \xi_2])^2\right] = \\ &= \mathrm{M}\left[((\xi_1 - \mathrm{M}[\xi_1]) + (\xi_2 - \mathrm{M}[\xi_2]))^2\right] = \\ &= \mathrm{M}\big[(\xi_1 - \mathrm{M}[\xi_1])^2 + (\xi_2 - \mathrm{M}[\xi_2])^2 + \\ &\quad + 2(\xi_1 - \mathrm{M}[\xi_1])(\xi_2 - \mathrm{M}[\xi_2])\big] = \\ &= \mathrm{M}\big[(\xi_1 - \mathrm{M}[\xi_1])^2\big] + \mathrm{M}\big[(\xi_2 - \mathrm{M}[\xi_2])^2\big] + \\ &\quad + 2\mathrm{M}\big[(\xi_1 - \mathrm{M}[\xi_1])(\xi_2 - \mathrm{M}[\xi_2])\big] = \\ &= \mathrm{D}[\xi_1] + \mathrm{D}[\xi_2] + 2\,\mathrm{cov}(\xi_1, \xi_2).\end{aligned}$$

Формула (6.44) узагальнюється на випадок n-вимірного випадкового вектора.

Справедлива наступна теорема.

Теорема 6.2. *Нехай* $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ — *випадковий вектор, для компонент якого існують* $k_{ij}\operatorname{cov}(\xi_i, \xi_j)$, $i, j = 1 \div n$. *Тоді для довільних сталих* $\lambda_i \in \mathbb{R}^1$, $i = 1 \div n$

$$\mathrm{D}\left[\sum_{i=1}^{n} \lambda_i \xi_i\right] = \sum_{i,j=1}^{n} k_{ij} \lambda_i \lambda_j. \tag{6.46}$$

Доведення. Нехай

$$\eta_n = \sum_{i=1}^{n} \lambda_i \xi_i$$

— випадкова величина, для якої справедлива рівність

$$\eta_n - \mathrm{M}[\eta_n] = \sum_{i=1}^{n} \lambda_i \big(\xi_i - \mathrm{M}[\xi_i]\big). \tag{6.47}$$

Дійсно, обчислимо

$$\mathrm{M}[\eta_n] = \mathrm{M}\left[\sum_{i=1}^{n} \lambda_i \xi_i\right] = \sum_{i=1}^{n} \lambda_i \mathrm{M}[\xi_i],$$

а тому рівність (6.47) є очевидною.

Так як

$$\big(\eta_n - \mathrm{M}[\eta_n]\big)^2 = \sum_{i,j=1}^{n} \lambda_i \lambda_j (\xi_i - \mathrm{M}[\xi_i])(\xi_j - \mathrm{M}[\xi_j]),$$

то

$$\begin{aligned}
\mathrm{M}\big[(\eta_n - \mathrm{M}[\eta_n])^2\big] &= \mathrm{D}\Big[\sum_{i=1}^{n} \lambda_i \xi_i\Big] = \\
&= \mathrm{M}\Big[\sum_{i,j=1}^{n} \lambda_i \lambda_j (\xi_i - \mathrm{M}[\xi_i])(\xi_j - \mathrm{M}[\xi_j])\Big] = \\
&= \sum_{i,j=1}^{n} \lambda_i \lambda_j \mathrm{M}\big[(\xi_i - \mathrm{M}[\xi_i])(\xi_j - \mathrm{M}[\xi_j])\big] = \\
&= \sum_{i,j=1}^{n} k_{ij} \lambda_i \lambda_j.
\end{aligned}$$

Тому

$$(\forall \lambda_j \in \mathbb{R})\colon \mathrm{D}\Big[\sum_{i=1}^{n} \lambda_i \xi_i\Big] = \sum_{i,j=1}^{n} k_{ij} \lambda_i \lambda_j.$$

■

При доведенні рівності (6.47) використано існування скінченного математичного сподівання $\mathrm{M}[\xi_j]$. Це гарантується існуванням $\operatorname{cov}(\xi_i, \xi_j)$. З доведеної теореми випливає важлива властивість дисперсійної матриці $\mathbb{K}$. Так як для всіх $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^1$ дисперсія $\mathrm{D}\Big[\sum\limits_{i=1}^{n} \lambda_i \xi_i\Big] \geqslant 0$, то і квадратична форма в правій частині (6.46) невід'ємно визначена. Необхідною та достатньою умовою невід'ємної визначеності квадратичної форми є невід'ємність всіх головних мінорів дисперсійної матриці $\mathbb{K} = \|k_{ij}\| = \|\operatorname{cov}(\xi_i, \xi_j)\|$. Це означає, що $\det \mathbb{K} = \det \|\operatorname{cov}(\xi_i, \xi_j)\|$, для довільних випадкових величин $\xi_1, \ldots, \xi_n$ задовільняє нерівність

$$\det \|\operatorname{cov}(\xi_i, \xi_j)\| \geqslant 0. \tag{6.48}$$

Для випадку $n = 2$ нерівність (6.48) дає

$$\begin{vmatrix} \mathrm{D}[\xi_1] & \operatorname{cov}(\xi_1, \xi_2) \\ \operatorname{cov}(\xi_1, \xi_2) & \mathrm{D}[\xi_2] \end{vmatrix} = \mathrm{D}[\xi_1]\mathrm{D}[\xi_2] - (\operatorname{cov}(\xi_1, \xi_2))^2 \geqslant 0$$

або

$$|\operatorname{cov}(\xi_1, \xi_2)| \leqslant \sqrt{\mathrm{D}[\xi_1]\mathrm{D}[\xi_2]}. \tag{6.49}$$

Для двовимірного випадкового вектора $\mathbf{x} = (\xi_1, \xi_2)$ вводять поняття *коефіцієнта кореляції.*

Означення 8. Коефіцієнтом кореляції між випадковими величинами ξ_1 та ξ_2 (для яких існують скінченні дисперсії $\mathrm{D}[\xi_1]$ та $\mathrm{D}[\xi_2]$) називають число

$$r = r(\xi_1, \xi_2) = \frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{\mathrm{D}[\xi_1]}\sqrt{\mathrm{D}[\xi_2]}}. \tag{6.50}$$

Розглянемо властивості коефіцієнта кореляції $r(\xi_1, \xi_2)$, означеного формулою (6.50).

1) Для довільних скалярних випадкових величини ξ_1 та ξ_2, що мають скінченні дисперсії ($\exists \mathrm{D}[\xi_1] < +\infty$, $\exists \mathrm{D}[\xi_2] < +\infty$),

$$|r(\xi_1, \xi_2)| \leqslant 1.$$

Це випливає з нерівності (6.49) та означення (6.50).

2) Якщо ξ_1 та ξ_2 — довільні *незалежні* випадкові величини зі скінченними дисперсіями ($\exists \mathrm{D}[\xi_1] < +\infty$, $\exists \mathrm{D}[\xi_2] < +\infty$), то

$$r(\xi_1, \xi_2) = 0.$$

Покажемо, що для *незалежних* випадкових величин ξ_1 та ξ_2

$$r(\xi_1, \xi_2) = 0.$$

З означення (6.44) та властивостей математичного сподівання випливає, що

$$\begin{aligned}
\operatorname{cov}(\xi_1, \xi_2) &= \mathrm{M}\big[(\xi_1 - \mathrm{M}[\xi_1])(\xi_2 - \mathrm{M}[\xi_2])\big] = \\
&= \mathrm{M}\big[(\xi_1\xi_2 - \xi_2\mathrm{M}[\xi_1] - \xi_1\mathrm{M}[\xi_2] + \mathrm{M}[\xi_1]\mathrm{M}[\xi_2])\big] = \\
&= \mathrm{M}[\xi_1\xi_2] - \mathrm{M}[\xi_2]\mathrm{M}[\xi_1] - \mathrm{M}[\xi_1]\mathrm{M}[\xi_2] + \mathrm{M}[\xi_1]\mathrm{M}[\xi_2] = \\
&= \mathrm{M}[\xi_1]\mathrm{M}[\xi_2] - \mathrm{M}[\xi_1]\mathrm{M}[\xi_2] = 0
\end{aligned}$$

і тому $r(\xi_1, \xi_2) = 0$.

Рівність нулю коефіцієнта кореляції є необхідною, але не достатньою умовою незалежності випадкових величин. З рівності $r(\xi_1, \xi_2) = 0$ не випливає незалежність випадкових величин ξ_1 та ξ_2.

Наведемо приклад таких величин.

Нехай ξ_1 — рівномірно розподілена на відрізку $[-1, 1]$ випадкова величина. Розглянемо випадкову величину $\xi_2 = \xi_1^2$. Так як $\mathrm{M}[\xi_1] = 0$, $\mathrm{M}[\xi_1^2] = \frac{1}{\xi}$, то

$$\begin{aligned}\operatorname{cov}(\xi_1, \xi_2) &= \mathrm{M}\big[\xi_1(\xi_1^2 - \mathrm{M}[\xi_1^2])\big] = \\ &= \mathrm{M}\big[\xi_1^3 - \xi_1 \mathrm{M}[\xi_1^2]\big] = \\ &= \mathrm{M}\big[\xi_1^3\big] - \mathrm{M}\big[\xi_1\big]\mathrm{M}\big[\xi_1^2\big] = \mathrm{M}\big[\xi_1^3\big].\end{aligned}$$

Але

$$\mathrm{M}\big[\xi_1^3\big] = \frac{1}{2}\int\limits_{-1}^{1} x^3\,\mathrm{d}x = 0.$$

Отже, $r(\xi_1, \xi_2) = 0$, хоч випадкові величини ξ_1 та $\xi_2 = \xi_1^2$ пов'язані функціональною залежністю.

Випадкові величини ξ_1 та ξ_2, для яких $r(\xi_1, \xi_2) = 0$ називають *некорельованими*. Приклад, розглянутий вище, показує, що з некорельованості двох випадкових величин взагалі кажучи не випливає їх незалежність. Однак, якщо випадковий вектор (ξ_1, ξ_2) розподілений нормально, то з некорельованості ξ_1 та ξ_2 випливає, що ξ_1 та ξ_2 — незалежні.

Доведення. Дійсно, нехай $\mathbf{x} = (\xi_1, \xi_2)$ розподілений нормально, тобто

$$\rho_{\mathbf{x}}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2}\frac{1}{\sqrt{1-p^2}}\exp\left\{-\frac{1}{2(1-p^2)}\left[\frac{(x_1-a)^2}{\sigma_1^2} - \right.\right. \\ \left.\left. - 2p\frac{(x_1-a)(x_2-b)}{\sigma_1\sigma_2} + \frac{(x_2-b)^2}{\sigma_2^2}\right]\right\}.$$

Тут $a = \mathrm{M}[\xi_1]$, $b = \mathrm{M}[\xi_2]$, $\mathrm{D}[\xi_1] = \sigma_1^2$, $\mathrm{D}[\xi_2] = \sigma_2^2$.

Обчислимо

$$
\begin{aligned}
\operatorname{cov}(\xi_1,\xi_2) &= \iint\limits_{-\infty}^{\infty}(x_1-a)(x_2-b)\rho_{\mathbf{x}}(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\frac{1}{\sqrt{1-p^2}}\iint\limits_{-\infty}^{\infty}(x_1-a)(x_2-b)\times \\
&\quad \times \exp\left\{-\frac{1}{2(1-p^2)}\left[\frac{(x_1-a)^2}{\sigma_1^2}-2p\frac{(x_1-a)(x_2-b)}{\sigma_1\sigma_2}+\right.\right. \\
&\qquad \left.\left.+\frac{(x_2-b)^2}{\sigma_2^2}\right]\right\}\mathrm{d}x_1\,\mathrm{d}x_2 = \\
&= \left|\begin{array}{l}\text{виділимо в показнику експоненти пов-}\\ \text{ний квадрат і зробимо підстановку}\\ t=\dfrac{1}{\sqrt{2(1-p^2)}}\left(\dfrac{x_2-b}{\sigma_2}-p\dfrac{x_1-a}{\sigma_1}\right)\end{array}\right| = \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\frac{1}{\sqrt{1-p^2}}\int\limits_{-\infty}^{\infty}(x_1-a)\exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2}\right\}\mathrm{d}x_1\times \\
&\times\sigma_2\sqrt{2(1-p^2)}\int\limits_{-\infty}^{\infty}\left[t+\frac{p\sigma_2^2}{\sigma_1}\sqrt{2(1-p^2)}(x_1-a)\right]\mathrm{e}^{-t^2}\mathrm{d}t = \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\frac{1}{\sqrt{1-p^2}}\frac{p\sigma_2^2}{\sigma_1}\sqrt{2(1-p^2)}\times \\
&\quad\times\int\limits_{-\infty}^{\infty}(x_1-a)^2\exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2}\right\}\mathrm{d}x_1 = \\
&= \frac{p\,\sigma_2}{\sigma_1}\mathrm{D}[\xi_1] = \frac{p\,\sigma_2}{\sigma_1^2}\sigma_1^2 = p\,\sigma_1\sigma_2,
\end{aligned}
$$

тому

$$
r(\xi_1,\xi_2)=\frac{\operatorname{cov}(\xi_1,\xi_2)}{\sqrt{\mathrm{D}[\xi_1]\mathrm{D}[\xi_2]}}=\frac{p\,\sigma_1\sigma_2}{\sigma_1\sigma_2}=p.
$$

Отже, параметр p — це коефіцієнт кореляції компонент ξ_1 та ξ_2 нормально розподіленого вектора $\mathbf{x} = (\xi_1, \xi_2)$. Якщо $p = 0$, то сукупна щільність нормально розподіленого вектора $\mathbf{x}$ є такою

$$\begin{aligned}
\rho_{\mathbf{x}}(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2} - \frac{(x_2-b)^2}{2\sigma_2^2}\right\} = \\
&= \frac{1}{2\pi\sigma_1} \exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2}\right\} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(x_2-b)^2}{2\sigma_2^2}\right\} = \\
&= \rho_{\xi_1}(x_1)\rho_{\xi_2}(x_2),
\end{aligned}$$

тобто ξ_1 та ξ_2 — незалежні. ■

Отже, *для нормально розподілених випадкових величин корельованість рівносильна до незалежності.*

Якщо ж компоненти випадкового вектора $\mathbf{x} = (\xi_1, \dots, \xi_n)$ є попарно корельованими, то в деяких випадках, знаючи коефіцієнт кореляції, можна відновити функціональну залежність між компонентами ξ_i, $i = 1 \div n$.

Справедлива наступна теорема.

Теорема 6.3. *Випадкові величини ξ_1 та ξ_2 пов'язані лінійною залежністю $\xi_2 = a\xi_1 + b$, $a, b \in \mathbb{R}^1$, тоді і тільки тоді, коли $|r(\xi_1, \xi_2)| = 1$, причому*

$$r(\xi_1, \xi_2) = \operatorname{sign} a = \begin{cases} 1, & a > 0, \\ -1, & a <> 0. \end{cases}$$

Доведення. *Необхідність.*

Нехай $\xi_2 = a\xi_1 + b$, $a, b \in \mathbb{R}^1$. Покажемо, що $|r(\xi_1, \xi_2)| = 1$. Покладемо $\mathrm{M}[\xi_1] = m$ та $\mathrm{D}[\xi_1] = \sigma_1^2$. Тоді

$$\mathrm{M}[\xi_2] = a\,m + b, \quad \mathrm{D}[\xi_2] = a^2\sigma_1^2.$$

Обчислимо

$$\begin{aligned}\mathrm{cov}(\xi_1, \xi_2) &= \mathrm{M}\big[(\xi_1 - m)(\xi_2 - \mathrm{M}[\xi_2])\big] = \\ &= \mathrm{M}\big[(\xi_1 - m)(a\,\xi_1 - a\,m)\big] = \\ &= \mathrm{M}\big[a\,\xi_1^2 - m\,a\,\xi_1 - a\,m\,\xi_1 + a\,m^2\big] = \\ &= a\left(\mathrm{M}[\xi^2] - \mathrm{M}^2[\xi_1]\right) = \\ &= a\,\mathrm{D}[\xi_1] = a\,\sigma_1^2.\end{aligned}$$

Маємо

$$r(\xi_1, \xi_2) = \frac{a\,\sigma_1^2}{\sqrt{a^2\sigma_1^2\sigma_2^1}} = \frac{a}{|a|} = \mathrm{sign}\,a.$$

Достатність.

Нехай $|r(\xi_1, \xi_2)| = 1$. Покажемо, що $\xi_2 = a\,\xi_1 + b$.

Скориставшись означенням $r(\xi_1, \xi_2)$ (формула (6.50)), отримуємо, що $\mathrm{cov}^2(\xi_1, \xi_2) = \mathrm{D}[\xi_1]\mathrm{D}[\xi_2]$. Нехай $t \in \mathbb{R}^1$ — довільне дійсне число. Розглянемо випадкову величину

$$\eta = t(\xi_1 - \mathrm{M}[\xi_1]) - (\xi_2 - \mathrm{M}[\xi_2]), \tag{6.51}$$

для якої завжди виконується нерівність

$$\mathrm{M}[\eta^2] \geqslant 0.$$

Так як

$$\mathrm{M}[\eta^2] = t^2\mathrm{D}[\xi] - 2t\,\mathrm{cov}(\xi_1, \xi_2) + \mathrm{D}[\xi_2],$$

то отримуємо, що $\forall t \in \mathbb{R}^1$

$$t^2\mathrm{D}[\xi_1] - 2t\,\mathrm{cov}(\xi_1, \xi_2) + \mathrm{D}[\xi_2] \geqslant 0.$$

Дискримінант квадратного тричлена, що є у лівій частині, дорівнює

$$\cos^2(\xi_1, \xi_2) - \mathrm{D}[\xi_1]\mathrm{D}[\xi_2] = 0.$$

Тоді цей квадратний тричлен має двократно вироджений дійсний корінь

$$t_0 = \frac{\mathrm{cov}(\xi_1, \xi_2)}{\mathrm{D}[\xi_1]},$$

що для η (формула (6.51)) дає

$$\eta = t_0(\xi_1 - \mathrm{M}[\xi_1]) - (\xi_2 - \mathrm{M}[\xi_2]),$$

і

$$\mathrm{M}[\eta^2] = 0.$$

Тому

$$\begin{aligned} 0 = \mathrm{M}[\eta^2] = \mathrm{M}\left[(t_0(\xi_1 - \mathrm{M}[\xi_1]) - (\xi_2 - \mathrm{M}[\xi_2]))^2\right] = \\ = \mathrm{M}\left[(t_0\xi_1 - \xi_2 - \mathrm{M}[t_0\xi_1 - \xi_2])^2\right] = \mathrm{D}[t_0\xi_1 - \xi_2]. \end{aligned}$$

З властивостей дисперсії випадкової величини випливає, що якщо $\mathrm{D}[t_0\xi_1 - \xi_2] = 0$, то з ймовірністю одиниця (майже напевно)

$$t_0\xi_1 - \xi_2 = \mathrm{const} = -b.$$

Поклавши $t_0 = a$, отримуємо, що майже напевно

$$\xi_2 = a\,\xi_1 + b.$$

Більше того, так як t_0 та $\mathrm{cov}(\xi_1, \xi_2)$ мають однакові знаки, то $r(\xi_1, \xi_2) = \mathrm{sign}\, a$. ■

§5. Багатовимірний нормальний розподіл

Нехай $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ — n-вимірний випадковий вектор і $\mathbb{K} = \|\,\mathrm{cov}(\xi_i, \xi_j)\|$ — коваріантна матриця його компонент ξ_i, $i = 1 \div n$, а $\mathbf{m} = (\mathrm{M}[\xi_1], \ldots, \mathrm{M}[\xi_n])$ — вектор математичного сподівання.

Кажуть, що вектор $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ розподілений нормально, якщо сукупна щільність розподілу

$$\begin{aligned} \rho_{\mathbf{x}}(x_1, \ldots, x_n) = N(\mathbf{m}, \mathbb{K}) = \\ = \frac{1}{(2\pi)^{n/2}(\det\|\mathbb{K}\|)^{n/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}^{\mathrm{T}} - \mathbf{m}^{\mathrm{T}})\mathbb{K}^{-1}(\mathbf{x} - \mathbf{m})\right\}. \end{aligned} \tag{6.52}$$

Тут $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{x}^{\mathrm{T}}$ — транспонований вектор відносно до $\mathbf{x}$.

Формула (6.52) показує, що багатовимірний нормальний розподіл повністю визначається математичним сподіванням та коваріантною матрицею випадкового вектора, тобто моментами першого та другого порядків. Твердження, що випадковий вектор $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ має розподіл $N(\mathbf{m}, \mathbb{K})$ означає, що він розподілений нормально (сукупна щільність задається формулою (6.52)), а його математичне сподівання — вектор $\mathbf{m}$, коваріаційна матриця — матриця $\mathbb{K}$.

Формула (6.52) дозволяє легко обчислити, що

$$\int\limits_{\mathbb{R}^n} \exp\left\{-\frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbb{C}\mathrm{u}\right\} \mathrm{d}u_1 \ldots \mathrm{d}u_n = \sqrt{\frac{(2\pi)^n}{\det \mathbf{C}}}, \tag{6.53}$$

який випливає з умови нормування функції $\rho_{\mathbf{x}}(x_1, \ldots, x_n)$:

$$\int\limits_{\mathbb{R}^n} \rho_{\mathbf{x}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n = 1.$$

Справедливе наступне твердження.

Якщо випадковий вектор $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ розподілений нормально, то некорельованість його компонент еквівалентна до їх незалежності.

ЗМІСТ