Київський національний університет імені Тараса Шевченка

Теорія ймовірностей, математична статистика і випадкові процеси

Навчальний посібник
для студентів фізичного факультету

Київ
Виробничо-поліграфічний центр
"Київський університет"
2001

Єжов С.М. Теорія ймовірностей, математична статистика і випадкові процеси: Навчальний посібник.
К.: ВПЦ "Київський університет", 2001, - 140 с.

Рецензент Ребенко О.Л., д-р. фіз.-мат. наук, професор

Затверджено Радою
фізичного факультету
9 жовтня 2000 року

Навчальне видання

Навчальний посібник
з курсу теорії ймовірностей,
математичної статистики і випадкових процесів

для студентів фізичного факультету

Автор ЄЖОВ Станіслав Миколайович

Редактор Т. Мельник
Молодший редактор ....

Вступ

Теорія ймовірностей, випадкові процеси і математична статистика створюють великі розділи математики та її застосувань. Їхній розвиток нерозривно пов'язаний із загальним розвитком науки і техніки, де все більш позначається потреба давати відповідну ймовірнісну інтерпретацію різним явищам і процесам. Теорія ймовірностей і випадкових процесів пропонує різноманітні математичні моделі для типових випадкових явищ та їх еволюційного розвитку. У рамках цих моделей вивчає притаманні їм ймовірнісні закономірності, розробляє методи розв'язку таких важливих задач, як прогнозування, керування та інші. Математична статистика розв'язує задачі оцінювання окремих параметрів і структури в цілому тієї чи іншої ймовірнісної моделі за статистичними даними, дає методи перевірки різних гіпотез, рекомендує правила планування самого експерименту для отримання необхідних статистичних даних.

Математична теорія ймовірностей набуває практичної цінності і наочний зміст у зв'язку з такими дійсними або уявними дослідами як, наприклад, підкидання монети сто разів, підкидання гральних кубиків, гра в рулетку, спостереження тривалості життя радіоактивного ядра або життя людини, схрещення двох сортів рослин. Сюди ж відносяться такі явища, як стать новонародженого, наявність випадкових шумів у системах зв'язку, контроль якості промислової продукції, положення частинки при дифузії, кількість подвійних зірок на різних ділянках неба тощо. Всі ці явища характеризуються тим, що для них відсутня *детерміністична регулярність* (тобто спостереження за ними не завжди приводять до однакових результатів). У той же час вони мають деяку *статистичну регулярність*. Це проявляється, наприклад, у статистичній сталості частот.

Дійсно, розглянемо підкидання правильної монети. Зрозуміло, що заздалегідь неможливо абсолютно достовірно передбачити результат кожного підкидання. Результати окремих експериментів носять нерегулярний характер і здається, що це позбавляє нас можливості встановити які -небудь закономірності, що пов'язані з цим експериментом. Проте, якщо провести велику кількість незалежних підкидань, то можна помітити, що для правильної монети буде спостерігатися цілком визначена статистична регулярність, коли частота випадання "герба" буде близька до 0.5.

Статистична сталість частот робить цілком правдоподібною гіпотезу про можливість кількісної оцінки *випадковості* того чи іншого явища, що здійснюється у результаті експериментів. Виходячи з цього, теорія ймовірності постулює наявність у події A деякої числової характеристики $P(A)$, що називається ймовірністю цієї події, природна властивість якої повинна полягати в тому, що зі зростанням кількості незалежних випробувань (експериментів) частота появи події A має наближатися до $P(A)$.

Відносно до розглянутого прикладу це означає, що ймовірність події A, яка полягає у випадінні "герба" при підкиданні правильної монети, природно вважати рівною 0.5. Кількість подібних прикладів, у яких інтуїтивне уявлення про чисельне значення ймовірності тієї чи іншої події складається досить легко, можна без зусиль примножити. При цьому всі вони будуть носити схожий характер і супроводжуватися невизначеними поняттями типу "чесне" підкидання, "правильна" монета, "незалежність" тощо.

Теорія ймовірностей, як і будь-яка точна наука, стала такою лише тоді, коли було чітко сформульовано поняття ймовірнісної моделі, та утворено її аксіоматику. Виникнення теорії ймовірностей як науки та її розвиток пов'язані з такими іменами: Паскаль (1623-1662), Ферма (1601-1665), Гюйгенс (1629-1695), Я.Бернуллі (1654-1705), Муавр (1667-1754), Лаплас (1749-1827), Пуассон (1781-1840), Гаусс (1777-1855), Чебишев (1821-1824), Марков (1856-1922), Ляпунов (1857-1918), Бернштейн (1880-1968), Мізес (1883-1953), Борель (1871-1956), Колмогоров (1903-1987).

Розділ 1

Елементарна теорія ймовірностей

1.1. Простір елементарних подій

У загальній теоретико-ймовірнісній схемі для кожного експерименту із випадковим результатом повинні бути вказані всі можливі результати, які відповідають таким вимогам: у результаті експерименту неодмінно відбувається один і тільки із цих результатів. Для нас несуттєва реальна природа цих результатів, важливим є лише те, що їх кількість N *скінчена*. Кожний такий результат прийнято називати *елементарною подією* ω. Сукупність елементарних подій

$$\Omega = \{\omega_1, \dots, \omega_N\} \tag{1.1}$$

будемо називати *простором елементарних подій*. Елементарні події є точками простору елементарних подій.

Будь-який результат експерименту з випадковим результатом у теорії ймовірностей прийнято називати *подією*. Кожну подію A можна ототожнювати з підмножиною $A \subset \Omega$, яка складається із тих точок $\omega \in \Omega$, при яких відбувається A. З цієї причини у подальшому ми не будемо робити різниці між подією A та відповідною підмножиною $A \subset \Omega$.

1.2. Дослід із скінченою кількістю рівноймовірних результатів

Розглянемо декілька прикладів структури простору елементарних подій.

Приклад 1. *При однократному підкиданні монети простір елементарних подій складається з двох точок:* $\Omega = \{\Gamma, P\}$, *де* Γ *- "герб",* P *- "решта".*

Приклад 2. (Підкидання грального кубика.) *Простір елементарних подій складається із шести точок:* $\Omega = \{1, 2, 3, 4, 5, 6\}$. *У цьому просторі можна, наприклад, виділити підпростір (подію)* A, *який описує природно парних чисел, тобто* $A = \{2, 4, 6\} \subset \Omega$.

У всіх цих прикладах жодній з елементарних подій не можна віддати перевагу перед іншими, вважати її більш ймовірною, ніж інші, і якщо N - загальна кількість елементарних подій, то випадання вважати елементарні події рівноймовірними і кожній події приписувати однакову ймовірність, що дорівнює $1/N$. Визначена таким чином ймовірність на практиці оцінює *частоту* даної елементарної події у серії великої кількості незалежних повторень експерименту, якщо під частотою розуміти відношення числа появи даної елементарної події до загального числа повторень експерименту. Після цього можна обчислити ймовірність будь-якого результату експерименту. А саме, якщо $N(A)$ розмірність підпростору $A \subset \Omega$ (кількість елементів ω, з яких складається цей підпростір), який визначає подію A, то випадання визначити ймовірність $P(A)$ події A як відношення числа елементарних подій, з яких складається A, до числа всіх елементарних подій:

$$P(A) = \frac{N(A)}{N}\,. \tag{1.2}$$

Зокрема, ймовірність випадання парного числа при підкиданні грального кубика дорівнює 0.5.

1.3. Дослід із нескінченною кількістю елементарних подій

Для багатьох експериментів характерним є те, що можлива нескінченна множина (навіть континуум) елементарних подій. У таких випадках ймовірність зручно задавати за допомогою т. зв. *густини ймовірності*. При цьому від додавання ймовірностей елементарних подій переходять до інтегрування густини ймовірності по множині, що відповідає деякій події A, а ймовірність кожної елементарної події дорівнює нулеві.

Приклад 3. *Нехай на відрізку $[a, b]$ довжини $l = b - a$ довільно ставиться точка. Яка ймовірність того, що вона попаде на відрізок $[\alpha, \beta]$, що міститься в $[a, b]$?*

$\triangleleft$ Оскільки ймовірність попасти в $[\alpha, \beta]$ не залежить від того, де саме на $[a, b]$ розташований відрізок $[\alpha, \beta]$, то шукана ймовірність $P(A)$ дорівнює $\dfrac{\beta-\alpha}{b-a}$, тобто відношенню довжин відрізків $[\alpha, \beta]$ і $[a, b]$. Відповідь буде такою ж, якщо замість відрізка $[\alpha, \beta]$ вибрати будь-яку підмножину відрізка $[a, b]$ і для неї можна визначити довжину, що дорівнює $(\beta - \alpha)$.

Для цього приклада густина ймовірності $p(\cdot)$ визначається рівністю

$$p(x) = \frac{1}{l}; \quad x \in [a, b]\,, \tag{1.3}$$

причому

$$p(A) = \int\limits_{\alpha}^{\beta} p(x)dx = \int\limits_{\alpha}^{\beta} \frac{dx}{l} = \frac{\beta-\alpha}{l}\,; \qquad \int\limits_{a}^{b} p(x)dx = 1\,. \tag{1.4}$$

$\triangleright$

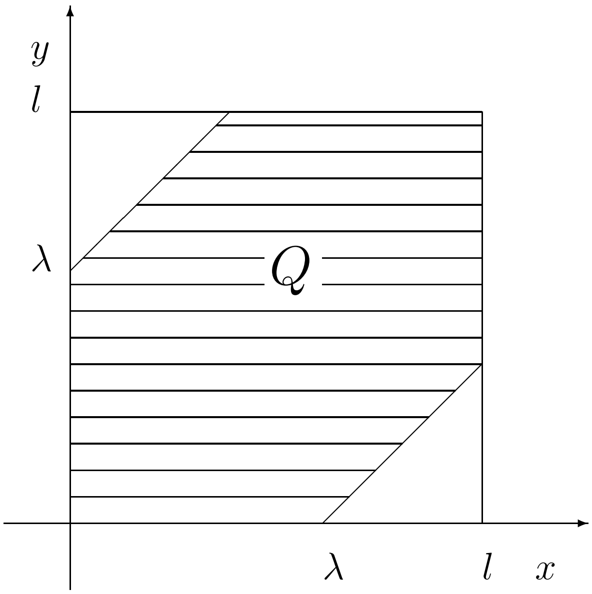

Рис. 1.1. До прикладу 4.

Приклад 4. *Нехай на відрізок $[a, b]$ довжини $l = b - a$ навмання ставляться дві точки. Яка ймовірність того, що відстань між ними буде не більша ніж $\lambda, 0 \le \lambda \le l$?*

$\lhd$ Зрозуміло, що задача еквівалентна такій: в квадрат $Q = \{(x, y) : 0 \le x \le l, 0 \le y \le l\}$ навмання кидається точка (x, y); яка ймовірність того, що точка попаде у заштриховану область на рис. 1.1? Тоді шукана ймовірність дорівнює відношенню площини заштрихованої області до площини квадрата Q:

$$p(A) = \int\limits_Q \frac{dxdy}{l^2} = \frac{(l^2 - (l - \lambda)^2)}{l^2} = \frac{2l\lambda - \lambda^2}{l^2}\,. \tag{1.5}$$

$\rhd$

1.4. Алгебра подій

Вже йшла мова про те, що з окремих елементарних подій ω можна складати деякі підмножини A, які мають назву *події*. Відштовхуючись від деякої заданої системи множин, що є подіями, можна утворювати нові події, що відповідають конструкціям висловів із логічними операціями "ні", "та", "або", які у теорії множин відповідають операціям "доповнення", "перетин", "об'єднання". Розглянемо визначення і властивості операцій над подіями, які характеризують алгебраїчну структуру будь-якої теоретико-ймовірнісної схеми.

1. Якщо подія A відбувається кожен раз, коли відбувається подія B, то будемо говорити, що подія A є *наслідком* події B, і записувати $B \subset A$ або $A \supset B$, тобто множина B є підмножиною A.

2. Якщо $A \subset B$ і $B \subset A$, то події A і B відбуваються або не відбуваються одночасно. У цьому випадку будемо писати $A = B$; множини A і B при цьому співпадають.

3. Подія, що полягає в тому, що *не* відбувається подія A, називається *протилежною* до події A і позначається $\overline{A}$. Множина $\overline{A}$ складається з точок множини Ω, що не належать до A, і називається *доповненням* множини A.

4. Якщо подія A не містить жодної елементарної події, то вона називається *неможливою* і позначається $\varnothing$. Очевидно, що протилежною $\varnothing$ є подія Ω, яка відбувається кожен раз і називається *достовірною*. Очевидно, що $\varnothing$ є пустою підмножиною Ω.

5. Подія C, що відбувається тоді і тільки тоді, коли відбуваються події A і B, називається *добутком* або *перетином* подій A і B, і позначається AB або $A \bigcap B$. Множина C складається із точок, що належать як множині A, так і множині B, і називається *перетином* множин A і B. Вона позначається $A \bigcap B$.

6. Події A і B називаються *несумісними*, якщо їх одночасне настання неможливе, тобто $A \bigcap B = \varnothing$. Несумісним подіям відповідають множини, що не перетинаються.

7. Подія C, що відповідає настанню хоча б однієї з подій A або B, називається *об'єднанням* або *сумою* подій A і B, і позначається $A \bigcup B$. У тому випадку, коли $A \bigcap B = \varnothing$, можна писати $C = A + B$. Відповідна множина C складається з тих точок Ω, які належать хоча б одній з множин A і B. *Об'єднання* або *сума* цих множин позначається $A \bigcup B$.

8. Подія C, що полягає в тому, що подія A відбувається, а подія B не відбувається, називається *різницею* подій A і B і позначається $A \setminus B$. У теорії множин множина $A \setminus B = A \bigcap \overline{B}$ складається з точок, що належать множині A і не належать множині B, і називається *різницею* множин A і B.

Розглянемо найпростіші властивості операцій над подіями. Перетин і об'єднання визначаються для довільної кількості подій. Операції об' єднання $\bigcup$ та перетину $\bigcap$ *асоціативні* та *комутативні* за означенням:

$$\begin{aligned} &(A \bigcup B) \bigcup C = A \bigcup (B \bigcup C), \quad A \bigcup B = B \bigcup A\,, \\ &(A \bigcap B) \bigcap C = A \bigcap (B \bigcap C), \quad A \bigcap B = B \bigcap A \end{aligned} \tag{1.6}$$

для будь-яких подій A, B і C.
За означенням

$$\overline{\overline{A}} = A, \quad \overline{A} = \Omega \setminus A, \quad A \setminus B = A \bigcap \overline{B} = \overline{B} \setminus \overline{A}\,. \tag{1.7}$$

Операції $\bigcup$ та $\bigcap$ *взаємно дистрибутивні* :

$$\begin{aligned} A \bigcup (B \bigcap C) &= (A \bigcup B) \bigcap (A \bigcup C)\,, \\ A \bigcap (B \bigcup C) &= (A \bigcap B) \bigcup (A \bigcap C)\,. \end{aligned} \tag{1.8}$$

Важливу роль у теорії ймовірностей відіграє *принцип двоїстості*, який може бути виражений такими співвідношеннями:

$$\overline{A \bigcup B} = \overline{A} \bigcap \overline{B}\,, \tag{1.9}$$

$$\overline{A \bigcap B} = \overline{A} \bigcup \overline{B}\,. \tag{1.10}$$

Доведемо формулу (1.9)

$$\omega \in \overline{A \bigcup B} \Leftrightarrow \omega \notin A \bigcup B \Leftrightarrow \omega \notin A, \omega \notin B \Leftrightarrow \omega \in \overline{A}, \omega \in \overline{B} \Leftrightarrow$$
$$\Leftrightarrow \omega \in \overline{A} \bigcap \overline{B}.$$

Принцип двоїстості справедливий для будь-якої кількості подій:

$$\overline{\bigcup_{\alpha\in S} A_\alpha} = \bigcap_{\alpha\in S} \overline{A}_\alpha, \quad \overline{\bigcap_{\alpha\in S} A_\alpha} = \bigcup_{\alpha\in S} \overline{A}_\alpha, \tag{1.11}$$

де символ $\bigcup_{\alpha\in S}$ ($\bigcap_{\alpha\in S}$) означає об'єднання (перетин) множини подій A_α, що відрізняються індексом $\alpha \in S$, який може також пробігати і незліченну множину значень.

До принципу двоїстості необхідно також віднести ще одне співвідношення

$$A \subset B \Leftrightarrow \overline{A} \supset \overline{B}. \tag{1.12}$$

Роль *принципу двоїстості* в теорії ймовірностей полягає в тому, що для будь-якого твердження, що відноситься до деякої системи подій, може бути сформульоване еквівалентне йому двоїсте твердження, в якому події повинні бути замінені на протилежні, операції об'єднання - на операції перетину і навпаки, та враховано співвідношення (1.12).

Розглянуті властивості операцій носять алгебраїчний характер. Таким чином, якщо є деяка система F множин $A \subseteq \Omega$, то за допомогою операцій $\bigcup, \bigcap, \setminus$ можна з елементів F побудувати нову систему множин, які також будуть подіями. Додаючи до цих подій *достовірну* Ω і *неможливу* $\varnothing$ події, ми отримаємо систему множин, яка називається *алгеброю*, тобто такою системою підмножин множини Ω, що:

1) якщо $A \in F$ і $B \in F$, то множини $A \bigcup B, A \bigcap B, A \setminus B$ також належать F;

2) разом із кожною подією A клас F містить протилежну подію $\overline{A}$.

З цих двох тверджень природна такі наслідки: якщо клас F не порожній, то звідси випливає, що $\Omega \in F$, оскільки $\Omega = A + \overline{A}$. Отже $\varnothing \in F$, оскільки $\varnothing = \overline{\Omega}$.

Конструкція алгебри подій дозволяє охарактеризувати множину усіх можливих результатів будь-якого експерименту з випадковими результатами, якщо множина Ω його елементарних результатів скінченна.

У випадку неперервної зміни можливих значень результату (кидання точки на відрізок) як систему подій в подальшому доведеться виділити спеціальний клас підмножин, більш широкий, ніж алгебра.

1.5. Класична теоретико-ймовірнісна модель

Ми вже зробили два кроки до визначення класичної ймовірнісної моделі експерименту зі скінченою кількістю елементарних результатів: виділили простір елементарних подій Ω і деяку систему його підмножин F, які утворюють алгебру і називаються подіями. В елементарній теорії ймовірностей як алгебра F звичайно береться алгебра *усіх* підмножин Ω. Як правило, виділення простору елементарних подій Ω і алгебри подій F не становить проблем.

Існує багато практичних ситуацій, коли симетрія приводить до рівноймовірності усіх елементарних результатів. У цьому випадку ми приходимо до *класичної ймовірнісної моделі.*

Нехай Ω - скінчений (або у більш загальному випадку нескінчений) простір елементарних подій деякого випадкового експерименту. Припустимо, що структура експерименту така, що на Ω можна вказати n елементарних подій $\omega_1, \ldots, \omega_n$, які мають такі властивості:
1) події ω_i, ω_j попарно несумісні, тобто ніякі дві з них не можуть відбутися одночасно

$$\omega_i \bigcap \omega_j = \varnothing, \quad i \neq j, \quad i, j = 1, \ldots, n\,; \tag{1.13}$$

2) $\omega_1, \ldots, \omega_n$ утворюють повну групу подій у тому розумінні, що при будь-якому результаті експерименту хоча б одна з них неодмінно відбувається, а це означає, що

$$\omega_1 + \ldots + \omega_n = \Omega\,; \tag{1.14}$$

3) події $\omega_1, \ldots, \omega_n$ рівноймовірні, або, іншими словами, жодній із них не можна віддати перевагу. Це *доматематична* вимога, і її не можна намагатися сформулювати у термінах яких-небудь більш елементарних властивостей.

У *класичній схемі* події $\omega_1, \ldots, \omega_n$, що відповідають умовам 1) - 3), називаються повною групою попарно несумісних рівноймовірних подій.

Допишемо кожній елементарній події $\omega_i \in \Omega$ деяку *вагу*, яку позначимо $p(\omega_i)$ і назвемо *ймовірністю* події ω_i. Ця ймовірність задовольняє таким умовам:
1) $0 \leq p(\omega_i) \leq 1$ (*невід'ємність*);
2) $p(\omega_1) + \ldots + p(\omega_n) = 1$ (*нормованість*).
Ймовірність $P(A)$ будь-якої події $A \in F$ має вигляд

$$P(A) = \sum_{\{i:\, \omega_i \in A\}} p(\omega_i)\,. \tag{1.15}$$

Тобто, в класичній схемі ймовірність визначається лише для тих результатів експерименту, які можуть бути представлені у вигляді об'єднання деяких із елементарних подій ω_i.
3) Внаслідок рівноймовірності усіх ω_i ймовірність події A визначається рівністю

$$P(A) = \frac{k}{n}\,, \tag{1.16}$$

де k - кількість доданків у (1.15). Це *класичне визначення ймовірності.*

Для того, щоб (1.16) вважалося коректним, достатньо показати єдиність розкладу

$$A = \omega_{i_1} + \omega_{i_2} + \ldots + \omega_{i_k}\,. \tag{1.17}$$

Для довільної події A відповідно до властивості дистрибутивності (1.8) маємо

$$A \bigcap \Omega = A \bigcap (\omega_1 + \ldots + \omega_n) = A \bigcap \omega_1 + \ldots + A \bigcap \omega_n\,. \tag{1.18}$$

Тоді, внаслідок властивості (1.13),

$$\begin{aligned} &A \bigcap \omega_j = \omega_{i_s},\ (j = i_s, \quad s = 1, \ldots, k)\,,\\ &A \bigcap \omega_j = \varnothing,\ (j \neq i_s)\,. \end{aligned} \tag{1.19}$$

Відповідно до свого класичного визначення знаходження ймовірності $P(A)$ звичайно зводиться до комбінаторних обчислень.

Приклад 5. (Задача про дні народження.) *Нехай в аудиторії є n студентів. Яка ймовірність того, що хоча б у двох збігаються дні народження?*

$\lhd$ Очевидно, що якщо $n \geq 366$, ця ймовірність дорівнює одиниці. Розглянемо випадки $n \leq 365$. Для групи з n студентів можливо $(365)^n$ комбінацій днів народження. Всі ці комбінації утворюють повну групу попарно несумісних рівноймовірних подій, причому ймовірність кожної комбінації дорівнює $1/(365)^n$. Кількість різних комбінацій, коли *жодна пара днів народження не збігіється*, дорівнює

$$365 \cdot 364 \cdots (365 - n + 1)\,. \tag{1.20}$$

Тоді ймовірність того, що на кожен день припадає *не більш одного дня народження*, дорівнює

$$P_n = \frac{365 \cdot 364 \cdots (365 - n + 1)}{(365)^n} = \frac{365!}{(365 - n)!(365)^n}\,. \tag{1.21}$$

Шукана ймовірність дорівнює $1 - P_n$. Відповідна таблиця значень має вигляд:

Таблиця 1.

n	*4*	*10*	*16*	*20*	*22*	*23*	*30*	*50*
$1-P_n$	*.016*	*.117*	*.284*	*.411*	*.476*	*.507*	*.706*	*.970*

Цікаво відмітити, що (всупереч очікуваному) розмір аудиторії, де з ймовірністю 1/2 знайдуться принаймні дві особи із співпадаючими днями народження, не такий вже великий: він дорівнює 23. $\rhd$

Приклад 6. (Система Максвелла-Больцмана.) *Нехай є r різних частинок, кожна з яких може знаходитись у будь-якій з n комірок (станів) незалежно від того, де при цьому знаходяться інші частинки. Скільки різних розміщень r частинок по n комірках?*

$\lhd$ Якщо усі розміщення (стани системи) рівноймовірні, то це статистика Максвелла-Больцмана. Усього існує n^r різних розміщень із ймовірностю кожного стану n^{-r}. $\rhd$

Приклад 7. (Статистика Бозе-Ейнштейна.) *Нехай є r однакових частинок, кожна з яких може знаходитись у будь-якій з n комірок (станів) незалежно від того, де при цьому знаходяться інші частинки. Скільки різних розміщень r частинок по n комірках?*

$\lhd$ Якщо усі розміщення (стани системи) рівноймовірні, то це статистика Бозе- Ейнштейна. Усього існує $N(n,r) = C_{n+r-1}^{n-1} = C_{n+r-1}^{r}$ різних розміщень (станів).
Дійсно, для однієї і двох частинок ця формула очевидна
$N(n,1) = n = C_n^1$ і $N(n,2) = n + n(n-1)/2 = C_n^1 + C_n^2 = C_{n+1}^2$. Тоді, якщо ця формула справедлива для r частинок, то для $r+1$ частинки маємо:
$N(n, r+1) = N(n,r) + N(n-1,r) + \ldots + N(1,r) = C_{n+r-1}^r + C_{n-1+r-1}^r + \ldots + C_r^r = (C_{n+r}^{r+1} - C_{n+r-1}^{r+1}) + (C_{n-1+r}^{r+1} - C_{n-1+r-1}^{r+1}) + \ldots + (C_{r+1}^{r+1} - C_r^r) + C_r^r = C_{n+r}^{r+1}$. $\rhd$

Приклад 8. (Статистика Фермі-Дірака.) *Нехай є r однакових частинок, кожна з яких може знаходитись в одній з n комірок (станів). При цьому в кожному стані може знаходитись не більш однієї частинки. Скільки різних розміщень r частинок по n комірках?*

$\lhd$ Якщо усі розміщення (стани системи) рівноймовірні, то це статистика Фермі- Дірака. Усього існує C_n^r різних розміщень (станів). $\rhd$

1.6. Властивості класичної ймовірності

У силу властивості (1.14) і рівноймовірності усіх елементарних подій для кожної з них маємо ймовірність

$$P(\omega_i) = \frac{1}{n}, \quad i = 1, \ldots, n. \tag{1.22}$$

Алгебра F подій у даному випадку складається із неможливої події $\varnothing$ і всіх можливих об'єднань одноточкових множин $\{\omega_i\}$, $i = 1, \ldots, n$, тобто усього з $C_n^0 + C_n^1 + \ldots + C_n^n = 2^n$ подій.

Для будь-якої події $A \in F$ ймовірність $P(A)$ визначається рівністю $P(A) = m/n$, де m - кількість елементарних подій, із яких складається A. (Очевидно, що ця властивість збігається з (1.16)).

Формально класична теоретико-ймовірнісна модель еквівалентна трійці (Ω, F, P), що складається із простору елементарних подій Ω із n точок, алгебри F, що складається із 2^n подій, і ймовірності $P(\cdot)$, визначеної для усіх подій із F.

Розглянемо властивості класичної ймовірності:

1) Для будь-якого $A \in F$ має місце $0 \leq P(A) \leq 1$ (адже $0 \leq m \leq n$).

2) Ймовірність достовірної події $A = \Omega$ дорівнює одиниці (адже для $A = \Omega$ маємо $m = n$). Ймовірність неможливої події $\varnothing$ дорівнює нулеві (оскільки для $A = \varnothing$ маємо $m = 0$).

3) Для несумісних подій A і B

$$P(A + B) = P(A) + P(B). \tag{1.23}$$

4) Для попарно протилежних подій A та $\overline{A}$

$$P(\overline{A}) = 1 - P(A). \tag{1.24}$$

$\lhd$ $A + \overline{A} = \Omega, \quad P(A + \overline{A}) = P(A) + P(\overline{A}) = 1.$ $\rhd$

5) Якщо подія A приводить до події B, тобто $A \subset B$, то

$$P(B \setminus A) = P(B) - P(A), \tag{1.25}$$

і $P(B) \geq P(A)$.

$\lhd$ $B = A + B \bigcap \overline{A}$, причому події A та $B \bigcup \overline{A}$ несумісні. Тоді внаслідок того, що $P(B) = P(A) + P(B \bigcap \overline{A})$, і в силу рівняння (1.23) маємо (1.25). $\rhd$

6) Для будь-яких двох подій A і B має місце

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B). \tag{1.26}$$

$\lhd$ $A \bigcup B = A + B \setminus A = A + B \bigcap \overline{A} = A + (B \bigcap \overline{A}) \bigcup (B \bigcap \overline{B}) = A + B \bigcap (\overline{A} \bigcup \overline{B}) = A + B \bigcap \overline{(A \bigcap B)} = A + B \setminus (A \bigcap B)$. Внаслідок того, що $A \bigcap (B \setminus A) = A \bigcap (B \bigcap \overline{A}) = \varnothing$, тобто події A та $B \setminus A$ несумісні, ми можемо використати (1.23) і (1.25): $P(A \bigcup B) = P(A + B \setminus A) = P(A) + P(B \setminus A) = P(A) + P(B \setminus (A \bigcap B)) = P(A) + P(B) - P(A \bigcap B)$. $\rhd$

7) Рівність (1.26) можна узагальнити на будь-яку кількість подій. Ймовірність того, що відбудеться хоча б одна із подій $A_1, \ldots, A_n$ дорівнює

$$\begin{aligned} P_{n,1} &= P(A_1 \bigcup A_2 \bigcup \cdots \bigcup A_n) = \sum_i P(A_i) - \\ &- \sum_{i<j} P(A_i \bigcap A_j) + \sum_{i<j<k} P(A_i \bigcap A_j \bigcap A_k) + \\ &+ \ldots + (-1)^{n-1} P(A_1 \bigcap A_2 \bigcap \ldots \bigcap A_n)\,. \end{aligned} \tag{1.27}$$

Наведемо більш загальні результати: ймовірність $Q_{n,m}$ того, що здійсниться рівно m подій із $A_1, \ldots, A_n$ дорівнює

$$\begin{aligned} Q_{n,m} &= S_m - C_{m+1}^1 S_{m+1} + C_{m+2}^2 S_{m+2} - \ldots + \\ &+ (-1)^{n-m} C_n^{n-m} S_n\,. \end{aligned} \tag{1.28}$$

Ймовірність $P_{n,m}$ того, що здійсниться не менш m подій із $A_1, \ldots, A_n$, дорівнює

$$\begin{aligned} P_{n,m} &= S_m - C_m^1 S_{m+1} + C_{m+1}^2 S_{m+2} - \ldots + \\ &+ (-1)^{n-m} C_{n-1}^{n-m} S_n\,. \end{aligned} \tag{1.29}$$

Тут

$$\begin{aligned} & S_j = \sum_{i_1<i_2<\ldots<i_j}^{n} P(A_{i_1} \bigcap A_{i_2} \bigcap \cdots \bigcap A_{i_j})\,, \\ & i_1, \ldots, i_j = 1, \ldots, n, \quad j = 1, \ldots, n, \quad S_0 = 1. \end{aligned} \tag{1.30}$$

Приклад 9. (Біноміальний розподіл.) *Припустимо, що монета підкидається n разів. Яка ймовірність того, що k разів з'явиться "герб"?*

$\lhd$ Результат спостережень можна записати у вигляді впорядкованого набору $(a_1, \ldots, a_n)$, де $a_i = 1$ при появі "герба" ("вдача") та $a_i = 0$ при появі "решти" ("невдача"). Простір усіх елементарних результатів має таку структуру

$$\Omega = \{\omega : \omega = (a_1, \ldots, a_n),\, a_i = 0, 1\}\,. \tag{1.31}$$

Кожній елементарній події ω можна приписати ймовірність

$$p(\omega) = p^{\sum_i a_i} q^{n - \sum_i a_i}\,, \tag{1.32}$$

де невід'ємні числа p та q такі, що $p + q = 1$.

Розглянемо усі події $\omega = (a_1, \ldots, a_n)$, для яких $\sum_i a_i = k$, тобто кількість випадання "герба" в серії n випробувань дорівнює k, $(k = 0, 1, \ldots, n)$. Згідно з прикладом 8. кількість таких елементарних подій дорівнює C_n^k, і для події A_k, що є об'єднанням таких подій,

$$A_k = \{\omega : \omega = (a_1, \ldots, a_n),\, a_1 + \ldots + a_n = k\},\ k = \overline{0, n}\,,$$

ймовірність буде дорівнювати

$$P(A_k) = C_n^k p^k q^{n-k}\,, \tag{1.33}$$

причому $\displaystyle\sum_{k=0}^{n} P(A_k) = 1$. Сукупність ймовірностей $\{P(A_0), \ldots,\ P(A_n)\}$ називається *біноміальним розподілом.*

Приклад 10. (Гіпергеометричний розподіл.) *Дана сукупність n об'єктів, серед яких k відмічених (наприклад, білетів, що виграли). Обирається навмання $n_1 \leq n$ об'єктів. Яка ймовірність того, що серед них виявиться k_1 відмічених?*

$\triangleleft$ Обрати n_1 об'єктів із n можна $C_n^{n_1}$ способами. Розмірність простору Ω дорівнює $C_n^{n_1}$. k_1 відмічених об'єктів із загальної їх кількості k можна обрати $C_k^{k_1}$ способами. Оскільки $k_1 \leq n_1$, то необхідно добрати ще $n_1 - k_1$ невідмічених об'єктів із їх загальної кількості $n - k$. Це можна зробити $C_{n-k}^{n_1-k_1}$ способами. Таким чином, кількість способів, що сприяють появі k_1 відмічених об'єктів серед n_1 обраних, дорівнює $C_k^{k_1} C_{n-k}^{n_1-k_1}$. У результаті, шукана ймовірність у відповідності до (1.16) дорівнює

$$P_{k,n}(k_1, n_1) = \frac{C_k^{k_1} C_{n-k}^{n_1-k_1}}{C_n^{n_1}}, \quad k_1 = 0, \ldots, \min(k, n_1) \tag{1.34}$$

Сукупність розподілів (1.34) називається *гіпергеометричним розподілом.*

За допомогою (1.34) підрахуємо ймовірності виграшу в спортлото 5 із 36:

$$P_{5,36}(k_1, 5) = \frac{C_5^{k_1} C_{36-5}^{5-k_1}}{C_{36}^5}\,,\ P(1,5) = 0.417,\ P(2,5) = 0.119\,,$$
$$P(3,5) = 0.0062\,,\ P(4,5) = 4.1 \cdot 10^{-4}\,,\ P(5,5) = 2.6 \cdot 10^{-6}\,.$$

$\triangleright$

Розділ 2

Аксіоматична побудова теорії ймовірностей

2.1. Система аксіом

До цього ми вважали всі елементарні події рівноймовірними. Проте завдання значень $p(\omega_i)$ насправді лежить поза межами теорії ймовірностей. Задачею теорії ймовірностей є не приписування тих чи інших значень $p(\omega_i)$, а обчислення ймовірностей складних подій (подій із F), відштовхуючись від ймовірностей елементарних подій. Для багатьох задач рівноймовірність ω_i не можна постулювати, і тому визначення ймовірностей у класичній схемі Лапласа за допомогою (1.16) стає неприйнятним. У рамках класичної моделі не можна описувати і інші задачі, наприклад геометричну ймовірність. У зв'язку з цим з'явилась необхідність переформулювати основні початкові положення теорії.

Нехай Ω - простір елементарних подій, F - алгебра подій (підмножин Ω). Наступні п'ять умов утворюють систему аксіом теорії ймовірностей.

1. Алгебра подій F називається σ-алгеброю, якщо для будь-якої послідовності подій $A_i \in F$, $i = 1, 2, \ldots$ їх об'єднання $A = A_1 \bigcup A_2 \bigcup \ldots = \bigcup_1^\infty A_i$ також належить F, тобто є подією.

Аксіома . F є σ-алгебра подій.

Підкреслимо, що мова йде лише про *зліченні* об'єднання і перетини.

2. На σ-алгебрі F визначається функція $P(\cdot)$, що приймає числові значення $P(A) \geq 0, A \in F$ і називається *ймовірністю.*

Ймовірність має такі властивості.

3. Для будь-яких подій A і B таких, що $A \bigcap B = \varnothing$, має місце *аксіома додавання ймовірностей*

$$P(A + B) = P(A) + P(B)\,. \tag{2.1}$$

Звідси випливає, що для будь-якого скінченного числа несумісних подій A_1,..,A_n має місце рівність

$$P(A_1 + \ldots + A_n) = P(A_1) + \ldots + P(A_n)\,. \tag{2.2}$$

4. Нехай події A_j, $j = 1, 2, \ldots,$ попарно несумісні: $A_i \bigcap A_j = \varnothing$, $i \neq j$, $i, j = 1, 2, \ldots$ і $A =$

$A_1 + A_2 + \ldots$. Тоді має місце *аксіома зліченної адитивності ймовірності*

$$P(A) = P(A_1) + P(A_2) + \ldots = \sum_{i=1}^{\infty} P(A_i)\,. \tag{2.3}$$

Відмітимо, що згідно з аксіомою 1 подія $A \in F$. Ця властивість також називається *аксіомою неперервності ймовірності.* Для цього розглянемо послідовність подій $B_1 = A_1$, $B_2 = A_1 + A_2, \ldots$. Подію A необхідно розуміти як границю послідовності $\{B_n\}$, $A = \lim\limits_{n\to\infty} B_n$. При цьому рівність (2.3) можна розуміти як властивість неперервності ймовірності

$$\begin{aligned} P(A) &= P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n) = \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} P(A_j) = \sum_{j=1}^{\infty} P(A_j)\,. \end{aligned} \tag{2.4}$$

5. Ймовірність достовірної події дорівнює одиниці

$$P(\Omega) = 1\,. \tag{2.5}$$

Простір елементарних подій Ω, σ-алгебра подій F та ймовірність $P(\cdot)$ на F, що задовольняють аксіомам теорії ймовірностей, утворюють т.зв. *ймовірнісний простір*, який ми будемо позначати (Ω, F, P).

Система аксіом теорії ймовірностей несуперечлива, оскільки існують (Ω,F,P), що задовольняють цим аксіомам, і неповна, оскільки ймовірність можна визначити багатьма способами у межах цих аксіом 2-5. Як приклад можна вказати на класичну теоретико-ймовірнісну модель, де Ω - скінчена множина, F - алгебра (і σ-алгебра) усіх підмножин Ω і ймовірність $P(\cdot)$ визначена для кожної підмножини $A \in F$ як відношення кількості точок, що утворюють A, до кількості усіх точок Ω.

2.2. Дискретні ймовірнісні простори

Ймовірнісний простір (Ω, F, P) називається *дискретним*, якщо множина $\Omega = \{\omega_1, \omega_2, \ldots\}$ скінченна або зліченна, F є σ- алгеброю усіх підмножин Ω (включаючи пусту множину $\varnothing$), ймовірність $P(\cdot)$ визначена для кожної одноточкової підмножини Ω:

$$P(\{\omega_j\}) = p_j \geq 0, \quad j = 1, 2, \ldots \quad \sum_{j=1}^{\infty} p_j = 1\,. \tag{2.6}$$

При цьому ймовірність будь-якої події $A \in F$ визначається рівністю

$$P(A) = \sum_{j:\omega_j \in A} p_j\,. \tag{2.7}$$

2.3. Властивості ймовірностей

Розглянемо властивості ймовірностей, які випливають із аксіом. Таким же чином, як і при доведенні властивостей класичної ймовірності знайдемо, що:
1)

$$P(\overline{A}) = 1 - P(A)\,, \tag{2.8}$$

оскільки $A + \overline{A} = \Omega$.
2) Якщо $A \subset B$, то

$$P(B \setminus A) = P(B) - P(A)\,, \tag{2.9}$$

оскільки $B = A + B \setminus A$. Отже, включення $A \subset B$ тягне за собою нерівність $P(A) \leq P(B)$ (монотонність ймовірності).
3) Для будь-яких подій $A_1, \ldots, A_n$ мають місце рівності (1.27), (1.28), (1.29). Доведення цих властивостей аналогічні класичним.
4) Нехай $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$ - послідовність подій, кожна з яких тягне за собою всі наступні. Якщо $A = \bigcup_1^\infty A_j$ - подія, яка полягає в тому, що відбувається хоча б одна з подій A_j, $(j = 1, 2, \ldots)$, то

$$P(A) = \lim_{n\to\infty} P(A_n)\,. \tag{2.10}$$

$\lhd$ Покладемо $A_0 = \varnothing$. Тоді

$$A = \bigcup_1^\infty A_j = (A_1 \setminus A_0) + (A_2 \setminus A_1) + \ldots + (A_n \setminus A_{n-1}) + \ldots \tag{2.11}$$

і

$$\begin{aligned} P(A) &= \sum_{j=1}^{\infty} P(A_j \setminus A_{j-1}) = \lim_{n\to\infty} \sum_{j=1}^{n} P(A_j \setminus A_{j-1}) = \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} (P(A_j) - P(A_{j-1})) = \lim_{n\to\infty} P(A_n)\,. \end{aligned} \tag{2.12}$$

$\rhd$
5) Якщо $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ - послідовність подій, кожна з яких тягне за собою усі попередні, і $A = \bigcap_1^\infty A_j$ - подія, яка полягає у тому, що відбуваються всі події $A_j, j = 1, 2, \ldots$, то

$$P(A) = \lim_{n\to\infty} P(A_n)\,. \tag{2.13}$$

$\lhd$ Відповідно до принципу двоїстості маємо $\overline{A_1} \subset \overline{A_2} \subset \ldots$ і $\overline{A} = \bigcup_1^\infty \overline{A_j}$. Тоді відповідно до (2.10) $P(\overline{A}) = \lim\limits_{n\to\infty} P(\overline{A_n})$ і отже

$$P(A) = 1 - P(\overline{A}) = \lim_{n\to\infty} (1 - P(\overline{A_n})) = \lim_{n\to\infty} P(A_n)\,. \tag{2.14}$$

$\triangleright$

Властивості 4) і 5) можна тлумачити, як *властивості неперервності ймовірностей* відносно монотонних граничних переходів. Дійсно, якщо $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$, то $A_n = \bigcup_1^n A_j$, і множину $A = \bigcup_1^\infty A_j$ природно назвати *границею монотонної послідовності* множин $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots : A = \lim\limits_{j\to\infty} A_j$. Тоді, відповідно до властивості 4), має місце

$$P(A) = P(\lim_{j\to\infty} A_j) = \lim_{j\to\infty} P(A_j)\,. \tag{2.15}$$

Таким же чином, якщо $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$, то $A_n = \bigcap_1^n A_j$, і множина $A = \bigcap_1^\infty A_j$ називається границею монотонної послідовності множин $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots : A = \lim\limits_{j\to\infty} A_j$. У даному випадку властивість 5) означає, що

$$P(A) = P(\lim_{j\to\infty} A_j) = \lim_{j\to\infty} P(A_j)\,. \tag{2.16}$$

2.4. Умовна ймовірність

Нехай заданий ймовірнісний простір (Ω, F, P). Розглянемо задачу визначення ймовірності події A, якщо відомо, що відбулася подія B, причому $P(B) > 0$. Тобто ми розглядаємо тільки ті елементарні події, які містяться в $A \bigcap B$. У зв'язку з цим подію B можна ототожнити із множиною Ω. Як наслідок, ми можемо зробити перехід $\Omega \to B$ і $F \to F_B$, де σ-алгебра подій F_B складається із подій вигляду $A_B = A \bigcap B$. Можна сказати, що на просторі елементарних подій B індукується σ-алгебра подій F_B. Вона індукується σ-алгеброю подій F.

Перевіримо, що F_B - σ-алгебра.

Нехай $A_B, C_B, C_B^j \in F_B$ $(A, C, C^j \in F)$. Тоді, використовуючи властивість дистрибутивності операцій об'єднання $\bigcup$ і перетину $\bigcap$, знайдемо

$$\begin{aligned}
&A_B \bigcup C_B = (A \bigcap B) \bigcup (C \bigcap B) = (A \bigcup C) \bigcap B = (A \bigcup C)_B\,,\\
&\bigcup_{j=1}^{\infty} C_B^j = \bigcup_{j=1}^{\infty} (C^j \bigcap B) = (\bigcup_{j=1}^{\infty} C^j) \bigcap B = (\bigcup_{j=1}^{\infty} C^j)_B\,,\\
&A_B \bigcap C_B = (A \bigcap B) \bigcap (C \bigcap B) = (A \bigcap C) \bigcap B = (A \bigcap C)_B\,.
\end{aligned}$$

Отже, $A_B \bigcup C_B\,,\ \bigcup_{j=1}^\infty C_B^j\,,\ A_B \bigcap C_B \in F_B$. Далі,

$$\overline{A_B} = B \setminus A_B = B \setminus A = B \bigcap \overline{A} = (\overline{A})_B\,, \tag{2.17}$$

тобто $\overline{A_B} \in F_B$, якщо $A_B \in F_B$.

Введемо на σ-алгебрі F_B ймовірність $P_B(\cdot)$:

$$P_B(A_B) = \frac{P(A \bigcap B)}{P(B)}. \tag{2.18}$$

Перевіримо, що $P_B(\cdot)$ відповідає аксіомам теорії ймовірностей.

Якщо $A_B \bigcap C_B = \varnothing$, то

$$\begin{aligned}
P_B(A_B + C_B) &= \frac{P((A \bigcap B) \bigcup (C \bigcap B))}{P(B)} = \frac{P(A_B) + P(C_B)}{P(B)} =\\
&= P_B(A_B) + P_B(C_B)\,,
\end{aligned}$$

тобто виконується аксіома додавання ймовірностей (2.1).

Аналогічно можна перевірити аксіому (2.3) (властивість неперервності ймовірності).

Очевидно, що $P_B(B)=\dfrac{P(B\bigcap B)}{P(B)}=1$, тобто виконується аксіома 5).

У випадку класичної ймовірності формула (2.18) має наглядне пояснення. У цьому випадку $\Omega=\{\omega_1,\dots,\omega_n\}$, причому події $\{\omega_1,\dots,\omega_n\}$ - рівноймовірні. Нехай B=$\{\omega_{j_1},\dots,\omega_{j_k}\}$, тобто $P(B)$=k/n. Нехай $A\bigcap B$=$\{\omega_{i_1},\dots,\omega_{i_s}\}\in\{\omega_{j_1},\dots,\omega_{j_k}\}$, тобто $P(A\bigcap B)=s/n$. Якщо B розглядати як новий простір елементарних подій, то ймовірність події A_B визначається як відношення числа елементарних подій s, що сприяють A_B, до загальної кількості елементарних подій k, тобто

$$P_B(A_B)=\frac{s}{k}=\frac{s/n}{k/n}=\frac{P(A\bigcap B)}{P(B)}\,. \tag{2.19}$$

Трійка (B,F_B,P_B) є новим ймовірнісним простором, що побудований у зв'язку з поставленою задачею. Але ймовірність $P_B(\cdot)$ можна розглядати і на σ-алгебрі F. На алгебрі F величина $P_B(\cdot)$ також є ймовірністю і позначається $P(\cdot|B)$

$$P(A|B)=\frac{P(A\bigcap B)}{P(B)}\,,\quad A\in F\,. \tag{2.20}$$

$P(A|B)$ як функція на F називається *умовною ймовірністю* події A при умові, що подія B відбулася.

Властивості $P(\cdot|B)$ аналогічні властивостям $P(\cdot)$:

1)

$$P(\Omega|B)=1\,. \tag{2.21}$$

2)

$$P(\overline{A}|B)=1-P(A|B)\,. \tag{2.22}$$

$\lhd$

$$\begin{aligned}
P(\overline{A}|B) &= \frac{P(\overline{A}\bigcap B)}{P(B)}=\frac{P(\overline{A\bigcup\overline{B}})}{P(B)}=\\
&= \frac{1-P(A\bigcup\overline{B})}{P(B)}=\frac{1-P(A)-P(\overline{B})+P(A\bigcap\overline{B})}{P(B)}=\\
&= \frac{P(B)-P(A)+P(A\bigcap\overline{B})}{P(B)}=1-\frac{P(A)-P(A\bigcap\overline{B})}{P(B)}=\\
&= 1-\frac{P(A\bigcap(B+\overline{B}))-P(A\bigcap\overline{B})}{P(B)}=1-\frac{P(A\bigcap B)}{P(B)}\,.
\end{aligned}$$

$\rhd$

3)

$$P(A_1+A_2|B)=P(A_1|B)+P(A_2|B). \tag{2.23}$$

$\triangleleft$

$$P(A_1 + A_2|B)=\frac{P((A_1+A_2)\bigcap B)}{P(B)}=\frac{P((A_1\bigcup B)\bigcup(A_2\bigcup B))}{P(B)}=$$
$$=\frac{P(A_1\bigcup B)}{P(B)}+\frac{P(A_2\bigcup B)}{P(B)}\,.$$

$\triangleright$

4)

$$P(A_1\bigcup A_2|B)=\frac{P(A_1\bigcap B)+P(A_2\bigcap B)-P((A_1\bigcap A_2)\bigcap B)}{P(B)}=$$
$$=P(A_1|B)+P(A_2|B)-P(A_1\bigcap A_2|B). \qquad (2.24)$$

$\triangleleft$ Має місце послідовність перетворень:

$(A_1\bigcup A_2)\bigcap B=(A_1\bigcap B)\bigcup(A_2\bigcap B)=$

$=(A_1\bigcap B)+(A_2\bigcap B)\setminus(A_1\bigcap B)=(A_1\bigcap B)+(A_2\bigcap B)\bigcap\overline{(A_1\bigcap B)}=$

$=(A_1\bigcap B)+\Big((A_2\bigcap B)\bigcap\overline{(A_1\bigcap B)}\Big)\bigcup(A_2\bigcap B)\bigcap\overline{(A_2\bigcap B)}=$

$=(A_1\bigcap B)+(A_2\bigcap B)\bigcap\Big(\overline{(A_1\bigcap B)}\bigcup\overline{(A_2\bigcap B)}\Big)=$

$=(A_1\bigcap B)+(A_2\bigcap B)\bigcap\overline{((A_1\bigcap B)\bigcap(A_2\bigcap B))}=$

$=(A_1\bigcap B)+(A_2\bigcap B)\setminus((A_1\bigcap B)\bigcap(A_2\bigcap B))\,.$

Очевидно, що

$(A_1\bigcap B)\bigcap((A_2\bigcap B)\setminus((A_1\bigcap B)\bigcap(A_2\bigcap B)))=$

$=(A_1\bigcap B)\bigcap\Big((A_2\bigcap B)\bigcap\overline{((A_1\bigcap B)\bigcap(A_2\bigcap B))}\Big)=\varnothing\,,$

тобто події $(A_1\bigcap B)$ та $(A_2\bigcap B)\setminus((A_1\bigcap B)\bigcap(A_2\bigcap B))$ несумісні. Тоді в силу (2.23) маємо

$$P((A_1\bigcup A_2)\bigcap B)=$$
$$=P(A_1\bigcap B)+P\left((A_2\bigcap B)\setminus((A_1\bigcap B)\bigcap(A_2\bigcap B))\right)$$

і використаємо (2.9). $\triangleright$

5) Очевидно, що має місце зліченна адитивність умовної ймовірності

$$P(A_1+A_2+\ldots|B)=\sum_{j=1}^{\infty}P(A_j|B)\,. \qquad (2.25)$$

$\triangleleft$ Дійсно,

$$\begin{aligned} P(A_1+A_2+\ldots|B) &= \frac{P((A_1+A_2+\ldots)\bigcap B)}{P(B)}= \\ &= \frac{P((A_1\bigcap B)+(A_2\bigcap B)+\ldots)}{P(B)}= \\ &= \frac{P(A_1\bigcap B)+P(A_2\bigcap B)+\ldots}{P(B)}\,. \end{aligned}$$

$\triangleright$

Приклад 11. *. Ймовірність аварії при запуску ракети дорівнює* 0.1*, у тому числі ймовірність аварії на старті дорівнює* 0.09*. Яка ймовірність аварії у випадку вдалого старту?*

$\lhd$ Нехай подія A - "аварія при запуску", $P(A) = 0.1$, подія B - "аварія на старті", $P(B) = 0.09$. Треба встановити $P(A|\overline{B})$. Очевидно, що відповідно до (2.22) маємо

$$P(A|\overline{B}) = 1 - P(\overline{A}|\overline{B}) = 1 - \frac{P(\overline{A}\bigcap\overline{B})}{P(\overline{B})} = \{B \subset A \quad \Rightarrow$$
$$\Rightarrow \quad \overline{B} \supset \overline{A}\} = 1 - \frac{P(\overline{A})}{P(\overline{B})} = 1 - \frac{1-P(A)}{1-P(B)} = \frac{1}{91} \approx 0.01092\,.$$

$\rhd$

З означення (2.20) випливає т.зв. *теорема добутку ймовірностей*

$$P(A\bigcap B) = P(A|B)P(B), \quad P(B) > 0. \tag{2.26}$$

Аналогічно маємо

$$\begin{aligned} &P(A\bigcap B\bigcap C) = P(A|B\bigcap C)P(B\bigcap C) = \\ &= P(A|B\bigcap C)P(B|C)P(C)\,,\ P(B\bigcap C),\, P(C) > 0\,. \end{aligned} \tag{2.27}$$

З математичного погляду результат є тривіальним, але його роль насправді нематематична. Ця теорема застосовується при визначенні ймовірності у тих випадках, коли за змістом задачі легко встановлюються умовні ймовірності.

Приклад 12. *Є* 30 *екзаменаційних білетів, серед яких є* 5 *"щасливих". Кому вигідніше тягнути білет - першому чи другому?*

$\lhd$ Нехай подія A - "перший витягує щасливий білет", B - "другий витягує щасливий білет", $\overline{A}$ - "перший не витягує щасливий білет", $\overline{B}$ - "другий не витягує щасливий білет". Зазначимо, що $A + \overline{A} = \Omega$, тоді $B = B\bigcap\Omega = B\bigcap A + B\bigcap\overline{A}$. Отже

$$\begin{aligned} P(B) &= P(B\bigcap A) + P(B\bigcap\overline{A}) = P(B|A)P(A) + \\ &+ P(B|\overline{A})P(\overline{A}) = \frac{4}{29}\cdot\frac{5}{30} + \frac{5}{29}\cdot\frac{25}{30} = \frac{1}{6} = P(A)\,. \end{aligned}$$

Додамо третього студента: подія C - "третій витягує щасливий білет". Очевидно, що

$$P(C|A\bigcap B) = \frac{3}{28}\,,\ P(C|A\bigcap\overline{B}) = \frac{4}{28}\,,$$
$$P(C|\overline{A}\bigcap B) = \frac{4}{28}\,,\ P(C|\overline{A}\bigcap\overline{B}) = \frac{5}{28}\,.$$

Тоді

$$\begin{aligned} &P(C){=}P(C\bigcap(A{+}\overline{A})\bigcap(B{+}\overline{B}){=}P\Big(\big((C\bigcap A){+}(C\bigcap\overline{A})\big)\bigcap(B{+}\overline{B})\Big){=} \\ &{=}P(C\bigcap A\bigcap B){+}P(C\bigcap A\bigcap\overline{B}){+}P(C\bigcap\overline{A}\bigcap B){+}P(C\bigcap\overline{A}\bigcap\overline{B}){=} \\ &{=}P(C|B\bigcap A)P(B|A)P(A){+}P(C|\overline{B}\bigcap A)P(\overline{B}|A)P(A){+} \\ &{+}P(C|B\bigcap\overline{A})P(B|\overline{A})P(\overline{A}){+}P(C|\overline{B}\bigcap\overline{A})P(\overline{B}|\overline{A})P(\overline{A}){=} \\ &{=}\frac{3}{28}{\cdot}\frac{4}{29}{\cdot}\frac{1}{6}{+}\frac{4}{28}{\cdot}\frac{25}{29}{\cdot}\frac{1}{6}{+}\frac{4}{28}{\cdot}\frac{5}{29}{\cdot}\frac{5}{6}{+}\frac{5}{28}{\cdot}\frac{24}{29}{\cdot}\frac{5}{6}{=}\frac{1}{6} = P(A) = P(B)\,. \end{aligned}$$

$\rhd$

2.5. Незалежність

Події A і B називаються *незалежними*, якщо

$$P(A\bigcap B) = P(A)P(B). \tag{2.28}$$

Якщо $P(B) > 0$, то згідно (2.26) та (2.28) маємо $P(A|B) = P(A)$, як і повинно бути за змістом умовної ймовірності. Аналогічно, якщо $P(A) > 0$, то $P(B|A) = P(B)$. У той же час визначення незалежності подій A і B на основі рівностей
$P(A|B) = P(A)$, $P(B|A) = P(B)$ не еквівалентно (2.28), оскільки в (2.28) не припускається існування умовних ймовірностей.

З означення незалежності випливає:
1) Ω і будь-яка подія A незалежні.
2) Будь-яка подія A і подія B незалежні, якщо $P(B) = 0$. Дійсно, із умови $A\bigcap B \subset B$ випливає, що $0 \leq P(A\bigcap B) \leq P(B) = 0$ і $P(A\bigcap B) = 0 = P(A)\cdot P(B)$.
3) Якщо A і B незалежні, то незалежні і такі пари подій: A і $\overline{B}$, $\overline{A}$ і B, $\overline{A}$ і $\overline{B}$. Дійсно, $A+\overline{A} = \Omega,\ A\bigcap B + \overline{A}\bigcap B = B,\ P(B) = P(A)\cdot P(B) + P(\overline{A}\bigcap B),\ P(\overline{A}\bigcap B) = (1-P(A))\cdot P(B) = P(\overline{A})\cdot P(B)$.
Аналогічно можна довести незалежність і для інших пар.

Розглянемо, яка конструкція подій визначає незалежність. Нехай є два дискретних простори

$$\Omega_1 = \left\{\omega_1^1,\dots,\omega_i^1,\dots\right\},\ \Omega_2 = \left\{\omega_1^2,\dots,\omega_j^2,\dots\right\}. \tag{2.29}$$

У ймовірнісному просторі (Ω_1, F_1, P_1) подіями є усі підмножини Ω_1 і ймовірність визначена для кожної елементарної події

$$P_1(\{\omega_i^1\}),\ i = 1,2,\dots \tag{2.30}$$

Аналогічно побудований (Ω_2, F_2, P_2). Визначимо (Ω, F, P), в якому Ω складається з усіх можливих впорядкованих пар $\omega_{ij} = (\omega_i^1\omega_j^2)$

$$\Omega = \left\{\omega_{ij} = (\omega_i^1,\omega_j^2),\ \omega_i^1 \subset \Omega_1,\ \omega_j^2 \subset \Omega_2\right\}. \tag{2.31}$$

Таке Ω називається *добутком* Ω_1 і Ω_1: $\Omega = \Omega_1 \times \Omega_2$. Відповідну σ-алгебру позначимо $F = F_1 \times F_2$. Визначимо ймовірність на F за допомогою рівності:

$$P(\{\omega_{ij}\}) = P(\{\omega_i^1\})P(\{\omega_j^2\}). \tag{2.32}$$

Ймовірнісний простір (Ω, F, P) називається *добутком просторів* (Ω_1, F_1, P_1) і $(\Omega_2,\ F_2, P_2)$:

$$(\Omega, F, P) = (\Omega_1, F_1, P_1) \times (\Omega_2, F_2, P_2). \tag{2.33}$$

При цьому для кожного $A \in F$ маємо

$$P(A) = \sum_{\omega_{ij}\in A} P(\{\omega_{ij}\}). \tag{2.34}$$

Розглянемо подію $A \in F$, яка складається із тих $\omega_{ij} = (\omega_i^1 \omega_j^2)$, в яких $\omega_i^1 \in A_1 \in F_1$, а ω_j^2 довільна: $\omega_j^2 \in \Omega_2$. Така подія називається *циліндричною* : $A = A_1 \times \Omega_2$. Відповідно до (2.32) і (2.34) маємо

$$P(A) = \sum_{\substack{\omega_i^1 \in A_1 \\ \omega_j^2 \in \Omega_2}} P_1(\{\omega_i^1\}) P_2(\{\omega_j^2\}) = P_1(A_1) P_2(\Omega_2) = P_1(A_1)\,. \tag{2.35}$$

Аналогічно для іншої циліндричної події $B = \Omega_1 \times A_2$ знайдемо $P(B) = P_2(A_2)$. Подія $A \bigcap B$ складається із тих пар $\omega_{ij} = (\omega_i^1, \omega_j^2)$, де $\omega_i^1 \in A_1$ і $\omega_j^2 \in A_2$. Тому

$$P(A \bigcap B) = \sum_{\substack{\omega_i^1 \in A_1 \\ \omega_j^2 \in A_2}} P_1(\{\omega_i^1\}) P_2(\{\omega_j^2\}) = P_1(A_1) P_2(A_2)\,. \tag{2.36}$$

Враховуючи (2.35) отримаємо

$$P(A \bigcap B) = P(A) P(B)\,, \tag{2.37}$$

тобто події A і B незалежні.

Приклад 13. (Гральні карти.)

$\lhd$ Нехай в колоді 52 карти. Розглянемо події: A_1 - витягнута дама, A_2 - витягнута карта пікової масті. Тоді

$$P(A_1) = \frac{4}{52} = \frac{1}{13},\ P(A_2) = \frac{13}{52} = \frac{1}{4},\ P(A_1) \cdot P(A_2) = \frac{1}{52}\,.$$

З іншого боку, ймовірність витягнути даму пікової масті дорівнює $P(A_1 \bigcap A_2) = \dfrac{1}{52}$. Тобто події A_1 і A_2 є незалежними. $\rhd$

Події $A_1, \ldots, A_n$ називаються *незалежними у сукупності*, якщо при будь-якому вибору подій $A_{i_1}, \ldots, A_{i_k}$ із даної сукупності виконується рівність

$$P(A_{i_1} \bigcap \ldots \bigcap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})\,. \tag{2.38}$$

З незалежності подій у сукупності випливає їх попарна незалежність. Обернене твердження є невірним, тобто із попарної незалежності подій не випливає їх незалежність у сукупності.

Приклад 14. (Різнокольоровий тетраедр.)

$\lhd$ Нехай три грані правильного тетраедра пофарбовані у червоний (r), синій (b) та зелений (g) кольори, а четверта грань пофарбована у три кольори відразу (rbg). Звідси випливає, що ймовірність впасти на грань, на якій х червоний колір, дорівнює: $P(r) = 1/2$. Аналогічно, $P(b) = P(g) = 1/2$. Умовна ймовірність того, що на грані є червоний колір при умові, що на ній вже є зелений, дорівнює

$$P(r|g) = \frac{P(r \bigcap g)}{P(g)} = \frac{1/4}{1/2} = \frac{1}{2}\,.$$

Таким чином, наприклад, $P(r\bigcap g) = 1/4 = P(r)\cdot P(g)$ і для інших пар аналогічно. Отже події r, b, g - попарно незалежні. Проте ймовірність впасти на грань із трьома кольорами $P(r\bigcap b\bigcap g) = 1/4 \neq 1/8 = P(r)\cdot P(b)\cdot (g)$, тобто події r, b, g не є незалежними у сукупності. $\triangleright$

Нехай (Ω, F, P) - ймовірнісний простір, $A_1,\ldots,A_n$ - повна група попарно несумісних подій

$$P(A_i\bigcap A_j) = \varnothing,\quad i\neq j,\quad A_1+\ldots+A_n=\Omega\,. \tag{2.39}$$

Якщо $B\in F$, то

$$B = B\bigcap A_1+\ldots+B\bigcap A_n\,. \tag{2.40}$$

Тому (див. (2.26))

$$P(B)=\sum_{j=1}^{n}P(B\bigcap A_j)=\sum_{j=1}^{n}P(B|A_j)P(A_j)\,. \tag{2.41}$$

Рівність (2.41) називається *формулою повної ймовірності.*

Якщо повна група складається із зліченної множини, то

$$B = B\bigcap A_1+\ldots+B\bigcap A_n+\ldots\,, \tag{2.42}$$

і внаслідок зліченної адитивності ймовірності випливає, що

$$P(B)=\sum_{j=1}^{\infty}P(B|A_j)P(A_j),\quad P(A_j)>0,\quad j=1,2,\ldots \tag{2.43}$$

Формула (2.43) називається *формулою повної ймовірності для зліченної множини.*

Якщо $P(B)>0$, то

$$P(A_k|B)=\frac{P(B|A_k)P(A_k)}{P(B)},\quad k=1,2,\ldots, \tag{2.44}$$

або

$$P(A_k|B)=\frac{P(B|A_k)P(A_k)}{\sum_j P(B|A_j)P(A_j)}\,. \tag{2.45}$$

Формули (2.44) та (2.45) називаються *формулами Байєса.*

Ці формули необхідно розуміти таким чином. У статистичних застосуваннях події $A_1,\ldots$, A_n, $(A_1+\ldots+A_n=\Omega)$ часто називають *гіпотезами*, а $P(A_i)$ - *апріорною ймовірністю* гіпотези. (*A priori* - до досліду, *A posteriori* - після досліду). Умовна ймовірність $P(A_j|B)$ розглядається як *апостеріорна ймовірність* гіпотези A_j після настання події B. Тобто відбуваються заміни

$$\Omega\to B\,,\ A_k\to A_k\bigcap B\,,\ P(A_k)\to\frac{P(A_k\bigcap B)}{P(B)}=P(A_k|B)\,. \tag{2.46}$$

Приклад 15. (Задача про розорення гравця.) *Нехай у результаті кожного туру гри капітал гравця змінюється на одну копійку* ± 1. *Гра закінчується при виконанні однієї із наступних умов: або гравець набирає капітал а копійок, або розорюється, тобто набирає* 0 *копійок. Знайти ймовірність розорення гравця.*

$\triangleleft$ Нехай $x(< 0)$ - початковий капітал гравця. Нехай $P(x)$ - ймовірність розорення. Парна і непарна кількості випадають у кожному турі гри з ймовірністю $1/2$. Нехай подія A - розорення гравця, подія A_1 - виграш у даному турі, подія A_2 - програш у даному турі. Тоді $A = A \bigcap A_1 + A \bigcap A_2$ і відповідно до (2.41) маємо

$$P(x){=}P(x|A_1)P(A_1){+}P(x|A_2)P(A_2){=}P(x+1)\cdot\frac{1}{2}+P(x-1)\cdot\frac{1}{2}\,.$$

Розв'язок цього різницевого рівняння шукаємо у вигляді ряду $P(x) = \sum\limits_{k=0}^{\infty} a_k x^k$. Тоді

$$\sum_{k=0}^{\infty} a_k x^k = \frac{1}{2}\sum_{k=0}^{\infty} a_k\left[(x+1)^k + (x-1)^k\right]\,.$$

Прирівнюючи коефіцієнти біля однакових степеней x, отримаємо

$$\begin{array}{lll} k=0 & a_0 = & a_0 + a_2 + a_4 + \ldots \\ k=1 & a_1 = & a_1 + 3a_3 + 5a_5 + \ldots \end{array}$$

Звідси $a_2 = a_3 = \ldots = 0$. Тобто $P(x) = a_0 + a_1 x$. Додаткові умови $P(0) = 1, \quad P(a) = 0$ дають однозначний розв'язок

$$P(x) = 1 - \frac{x}{a}\,.$$

$\triangleright$

Приклад 16. *В урні знаходяться дві монети:* A_1 *- симетрична монета з ймовірністю "герба", що дорівнює* $1/2$, *і* A_2 *- несиметрична монета з ймовірністю "герба", що дорівнює* $1/3$. *Навмання виймається і підкидається одна із монет. Припустимо, що випав "герб". Питання: яка ймовірність того, що обрана монета симетрична?*

$\triangleleft$ Побудуємо ймовірнісну модель. Елементарними подіями оберемо множину $\Omega = \{(A_1\Gamma), (A_1P), (A_2\Gamma), (A_2P)\}$, що описує усі можливі результати вибору і підкидання. Ймовірності $P(\omega)$ усіх елементарних подій повинні бути обрані таким чином, щоб задовольняти умовам задачі:

$$P(A_1) = P(A_2) = \frac{1}{2}\,, \quad P(\Gamma|A_1) = \frac{1}{2}\,, \quad P(\Gamma|A_2) = \frac{1}{3}\,.$$

Тоді ймовірності елементарних подій визначаються однозначно:

$$P(A_1\Gamma) = \frac{1}{4}\,, \quad P(A_1P) = \frac{1}{4}\,, \quad P(A_2\Gamma) = \frac{1}{6}\,, \quad P(A_1P) = \frac{1}{3}\,.$$

Згідно з формулою Байєса (2.45), шукана ймовірність дорівнює

$$P(A_1|\Gamma){=}\frac{P(\Gamma|A_1)P(A_1)}{P(\Gamma|A_1)P(A_1)+P(\Gamma|A_2)P(A_2)}{=}\frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{3}}{=}\frac{3}{5}\,.$$

Аналогічно $P(A_2|\Gamma) = \dfrac{2}{5}\,.$ $\triangleright$

Розділ 3

Послідовність незалежних випробувань

3.1. Схема Бернуллі

Розглянемо дискретний ймовірнісний простір(Ω,F,P), де ймовірність P визначена для кожної елементарної події $\omega_i \in \Omega$ за допомогою рівності $P(\{\omega_i\}) = p_i$, $i = 1, 2, \ldots$, $\sum_i p_i = 1$. Будемо вважати, що ймовірнісному простору $(\Omega, F,$
$P)$ можна співставити дослід S, а $\omega_1, \omega_2, \ldots$ - можливі елементарні результати цього досліду. Тоді досліду S, що був повторений двічі, можна співставити ймовірнісний простір $(\Omega_2, F_2,$-$P_2) = (\Omega, F, P) \times (\Omega, F, P)$. Зараз елементарними подіями будуть упорядковані пари подій $(\omega_i\omega_j) \in \Omega \times \Omega = \Omega_2$ і $F_2 = F \times F$ буде σ - алгеброю підмножин Ω_2. Ймовірність P_2 на F_2 можна визначити багатьма способами, але якщо результати першого досліду ніяким чином не впливають на результати другого, то відповідно до (2.32) необхідно покласти

$$P_2(\{\omega_i\omega_j\}) = p_i p_j, \quad i, j = 1, 2, \ldots \tag{3.1}$$

Очевидно, що n разів незалежно повтореному досліду S відповідає ймовірнісний простір $(\Omega_n, F_n, P_n) = [\times(\Omega, F, P)]^n$, де $\Omega_n = [\times\Omega]^n$, $F_n = [\times F]^n$, а ймовірність P_n задана рівностями

$$P_n(\{\omega_i\omega_j \ldots \omega_k\}) = p_i p_j \cdots p_k, \quad i, j, \ldots, k, = 1, 2, \ldots \tag{3.2}$$

Нехай (Ω, F, P) - дискретний ймовірнісний простір. *Послідовністю n незалежних випробувань* називається ймовірнісний простір (Ω_n, F_n, P_n), в якому елементарними подіями х послідовності $(\omega_{i_1} \ldots \omega_{i_n})$ і ймовірність визначена для кожної елементарної події за допомогою рівності (3.2). Послідовність незалежних випробувань називається *схемою Бернуллі*, якщо $\Omega = (\omega_1, \omega_2)$, тобто дослід S має лише два елементарних результати. Успіх: $P(\{\omega_1\}) = p$, невдача: $P(\{\omega_2\}) = q = 1 - p$.

У схемі Бернуллі простір Ω_n складається із 2^n елементів, причому

$$P_n(\{\omega_{i_1} \ldots \omega_{i_n}\}) = p^k q^{n-k}\,, \tag{3.3}$$

де k - кількість успіхів у послідовності $\omega_{i_1} \ldots \omega_{i_n}$ елементарних результатів.

Знайдемо ймовірність того, що в схемі Бернуллі в серії n випробувань успіх матиме місце рівно $k \leq n$ разів. Оскільки не має значення, коли саме в цих випробуваннях будуть спостерігатися ці k успіхів, то подія, яка складається з того, що успіх відбувся k разів, буде об'єднанням різних подій типу $(\omega_{i_1}, \ldots, \omega_{i_n})$, де ω_1 зустрічається k разів. Таких подій буде C_n^k і, оскільки всі вони несумісні, то

$$P_n(A_k) = p_n(k) = C_n^k p^k q^{n-k}\,. \tag{3.4}$$

Сукупність (3.4) називається *біноміальним розподілом.* Назва є наслідком того, що (3.4) є загальним членом розкладу бінома

$$1 = (p+q)^n = \sum_{k=0}^{n} = C_n^k p^k q^{n-k}\,. \tag{3.5}$$

Ця рівність показує, що елементарні події, які містять $k = 0, 1, \ldots, n$ успіхів, утворюють повну групу попарно несумісних подій.

Нехай $\Omega = \{\omega_1, \ldots, \omega_r\}$ і $p(\{\omega_i\}) = p_i, i = 1, \ldots, r$. Нехай відбудеться n випробувань. У результаті отримаємо елементарну подію $\{\omega_{i_1} \cdots \omega_{i_n}\}$ із ймовірністю $p_1^{s_1} \cdots p_r^{s_r}$, де $s_1, \ldots, s_r$ кількість елементарних подій $\omega_1, \ldots, \omega_r$, відповідно, в послідовності $\{\omega_{i_1} \cdots \omega_{i_n}\}$ і $s_1 + \ldots + s_r = n$. Тоді при n випробувань ймовірність того, що ω_1 спостерігається s_1 разів, ω_2 спостерігається s_2 разів і т.д., дорівнює

$$P_n(s_1, \ldots, s_r) = \frac{n!}{s_1! \ldots s_r!}\, p_1^{s_1} \cdots p_r^{s_r}, \quad s_1 + \ldots + s_r = n\,. \tag{3.6}$$

Це *поліноміальний (мультиноміальний) розподіл.* Вираз (3.6) є загальним членом розкладу

$$1 = (p_1 + \ldots + p_r)^n = \sum_{s_1 \ldots s_r} \frac{n!}{s_1! \ldots s_r!}\, p_1^{s_1} \cdots p_r^{s_r}\,. \tag{3.7}$$

Множник $\dfrac{n!}{s_1! \ldots s_r!}$ дорівнює кількості можливих елементарних подій у n випробуваннях, у яких ω_i спостерігається s_i разів. Дійсно, елементарну подію ω_1 (s_1 разів) можна розкласти по n місцях $C_n^{s_1}$ різними способами; елементарну подію ω_2 (s_2 разів) можна розкласти по $n - s_1$ місцях, що залишились, $C_{n-s_1}^{s_2}$ різними способами і т.д. У результаті отримаємо

$$\frac{n!}{s_1!(n-s_1)!} \cdot \frac{(n-s_1)!}{s_2!(n-s_1-s_2)!} \cdots \frac{(n-s_1-\ldots-s_{r-1})!}{s_r!(n-s_1-\ldots-s_r)!} = \frac{n!}{s_1! \ldots s_r!}\,.$$

Приклад 17. (Гра в бридж.) *Яка ймовірність того, що кожен гравець отримає по одному тузу?*

◁ В цій грі 52 карти розподіляються на 4 рівні групи і кількість різних розкладів дорівнює $52!/(13!)^4 \approx 5,36 \cdot 10^{28}$. Чотири тузи можна упорядкувати $4! = 24$ різними способами (при умові що кожен гравець отримає по одному тузу). 48 карт, що залишились, можна розподілити $48!/(12!)^4$ різними способами. Тоді ймовірність того, що кожний гравець отримає по тузу, дорівнює $\dfrac{24 \cdot \dfrac{48!}{(12!)^4}}{\dfrac{52!}{(13!)^4}} \approx 0,105\,.$ ▷

3.2. Розподіл Пуассона

Формули біноміального розподілу (3.4) при великих n приводять до громіздких обчислень. Тому будемо шукати наближені, але прості формули для обчислення відповідних ймовірностей. У багатьох застосуваннях можна зустріти ситуацію, коли n є відносно великим, p - відносно малим, а добуток

$$pn = \lambda \tag{3.8}$$

і не малий, і не великий. Знайдемо у цьому випадку наближений вираз для ймовірності (3.4).

Теорема 1. (Пуассона.) *Нехай дана послідовність $\{s_n\}$ серій незалежних випробувань, що складаються відповідно із $1, 2, \ldots, n, \ldots$ випробувань, і нехай ймовірність події A при кожному випробуванні n-ої серії дорівнює λ/n, де λ не залежить від n. Тоді ймовірність $P_n(m)$ того, що кількість настання події A в n-ій серії буде дорівнювати m, при $n \to \infty$ і фіксованому m прямує до $\dfrac{\lambda^m}{m!}e^{-\lambda}$.*

$\lhd$ Відповідно до (3.4) має місце співвідношення

$$\begin{aligned}
&\lim_{n\to\infty} P_n(m) = \lim_{n\to\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} = \\
&= \lim_{n\to\infty} \frac{n!}{m!(n-m)!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} = \\
&= \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \left\{ \frac{n(n-1)\cdots(n-m+1)}{n^m} \left(1 - \frac{\lambda}{n}\right)^{-m} \right\} = \\
&= \frac{\lambda^m}{m!} e^{-\lambda}\,.
\end{aligned}$$

$\rhd$

Розподіл

$$P(m) = \frac{\lambda^m}{m!} e^{-\lambda}, \quad m = 0, 1, \ldots, \quad \lambda > 0 \tag{3.9}$$

називається *класичним пуассонівським наближенням для біноміального розподілу (3.4)* або *розподілом Пуассона.*

На практиці формула (3.9) слугує добрим наближенням для (3.4), якщо $n \geq 100, \quad 0 \leq np \leq 10, \quad m = 0, 1, \ldots, n.$

Приклад 18. (Радіоактивний розпад.) *Ймовірність зареєструвати частинку дорівнює 10^{-4}. Яка найменша кількість частинок повинна вилетіти із джерела для того, щоб із ймовірністю не менше $0,99$ зареєструвати більше трьох частинок?*

$\lhd$ Нехай n - шукана кількість частинок. Подія A - лічильник зареєстрував більш трьох частинок. Тоді $P(A) = 1 - P(\overline{A})$ та

$$\begin{aligned}
&P(\overline{A}) = P_n(0) + P_n(1) + P_n(2) + P_n(3) \approx P(0) + P(1) + P(2) + P(3) = \\
&= e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\right) \leq 0.01\,.
\end{aligned}$$

Звідси $\lambda \approx 10,7$ та $m = \lambda/p \approx 10,7 \cdot 10^4$ частинок. $\rhd$

3.3. Локальна теорема Муавра-Лапласа

Для випадку $p \ll 1$ має місце розподіл Пуассона (3.9). Якщо ж p не прямує до нуля, то справедлива інша гранична формула при $0 < p < 1$.

Теорема 2. (Локальна теорема Муавра-Лапласа.) *Якщо ймовірність події A в n незалежних випробуваннях дорівнює p, $0 < p < 1$, то ймовірність $P_n(m)$ того, що у цих випробуваннях подія відбудеться m разів, задовольняє при $n \to \infty$ співвідношенню*

$$\lim_{n\to\infty} \frac{P_n(m)}{\dfrac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\dfrac{x_m^2}{2}\right]} = 1\,, \tag{3.10}$$

де $\sigma_n = \sqrt{npq}$, $x_m = \dfrac{m-np}{\sigma_n}$, $q = 1-p$, $x_m \in [a,b]$, $a < b$ - будь-які обмежені числа. Прямування до одиниці є рівномірним відносно усіх m, для яких $x_m \in [a,b]$.

$\triangleleft$ Відповідно до (3.4) має місце

$$\sqrt{npq}\,P_n(m) = \sqrt{npq}\,\frac{n!}{m!(n-m)!}\,p^m q^{n-m}\,. \tag{3.11}$$

Використаємо формулу Стірлінга

$$k! = \sqrt{2\pi k}\,k^k e^{-k} e^{\theta_k}, \quad |\theta_k| \le \frac{1}{12k}\,. \tag{3.12}$$

Тоді

$$\begin{aligned} &\sqrt{npq}P_n(m) = \sqrt{npq}\cdot\frac{1}{2\pi}\cdot\frac{n^n e^{\theta_n-\theta_m-\theta_{n-m}}}{m^m(n-m)^{n-m}}\cdot\sqrt{\frac{n}{m(n-m)}}\cdot \\ &\cdot p^m q^{n-m} = \frac{1}{\sqrt{2\pi}}\left\{\left(\frac{np}{m}\right)^m\left(\frac{nq}{n-m}\right)^{n-m}\right\}\cdot\left\{\sqrt{\frac{n^2pq}{m(n-m)}}\right\}\cdot \\ &\cdot\left\{e^{\theta_n-\theta_m-\theta_{n-m}}\right\} = \frac{1}{\sqrt{2\pi}}A_n(x_m)B_n(x_m)C_n(x_m)\,. \end{aligned} \tag{3.13}$$

Нехай $[a,b]$ - довільний обмежений інтервал; будемо розглядати такі m, для яких $x_m \in [a,b]$. Оскільки $x_m = \frac{m-np}{\sigma_n}$, то

$$\begin{aligned} &m = np + x_m\sqrt{npq}\,, \\ &n - m = n(1-p) - x_m\sqrt{npq} = nq - x_m\sqrt{npq}\,. \end{aligned} \tag{3.14}$$

Розглянемо граничне значення множника $C_n(x_m) = \exp(\theta_n - \theta_m - \theta_{n-m}) = \exp(\theta)$:

$$\begin{aligned} |\theta| &\le |\theta_n| + |\theta_m| + |\theta_{n-m}| \le \\ &\le \frac{1}{12}\left(\frac{1}{n} + \frac{1}{np+x_m\sqrt{npq}} + \frac{1}{nq-x_m\sqrt{npq}}\right) = \\ &= \frac{1}{12n}\left(1 + \frac{1}{p + x_m\sqrt{\dfrac{pq}{n}}} + \frac{1}{q - x_m\sqrt{\dfrac{pq}{n}}}\right). \end{aligned} \tag{3.15}$$

Звідси випливає, що за ознакою Вейєрштрасса, $\theta \to 0$ при $n \to \infty$ рівномірно по $x_m \in [a, b]$.

Із явного вигляду

$$B_n(x_m)=\sqrt{\frac{n^2pq}{m(n-m)}}=\sqrt{\frac{1}{(1+x_m\sqrt{\frac{q}{np}})(1-x_m\sqrt{\frac{p}{nq}})}}\underset{n\to\infty}{\Rightarrow} 1 \tag{3.16}$$

видно, що це граничне значення досягається рівномірно по $x_m \in [a, b]$.

Розглянемо граничне значення $A_n(x_m)$. Для цього використаємо формулу

$$\ln(1+z) = z - \frac{z^2}{2} + O(z^3)\,,\ |z| < 1\,. \tag{3.17}$$

Тоді

$$\begin{aligned}
&\ln A_n(x_m) = -m\ln\left(\frac{m}{np}\right) - (n-m)\ln\left(\frac{n-m}{nq}\right) = \\
&= -(np + x_m\sqrt{npq})\ln\left(1 + x_m\sqrt{\frac{q}{np}}\right) - \\
&-(nq - x_m\sqrt{npq})\ln\left(1 - x_m\sqrt{\frac{p}{nq}}\right) = \\
&= -(np + x_m\sqrt{npq})\left(x_m\sqrt{\frac{q}{np}} - x_m^2\frac{q}{2np} + O(n^{-3/2})\right) - \\
&-(nq - x_m\sqrt{npq})\left(-x_m\sqrt{\frac{p}{nq}} - x_m^2\frac{p}{2nq} + O(n^{-3/2})\right) = \\
&= -\Big[x_m\sqrt{npq} + x_m^2 q - x_m^2\frac{q}{2} - x_m^3 O(n^{-1/2}) + O(n^{-1/2}) - \\
&-x_m\sqrt{npq} + x_m^2 p - x_m^2\frac{p}{2} + x_m^3 O(n^{-1/2}) + O(n^{-1/2})\Big] = \\
&= -\frac{1}{2}x_m^2 + O(n^{-1/2})\,.
\end{aligned} \tag{3.18}$$

Причому при $n \to \infty$ оцінку $O(n^{-1/2})$ можна вважати незалежною від m. З (3.13), (3.15), (3.16), (3.18) випливає (3.10).$\triangleright$

Ця теорема дає оцінку при великих n

$$P_n(m) \approx \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{(m-np)^2}{2\sigma_n^2}\right)\,. \tag{3.19}$$

3.4. Інтегральна теорема Муавра-Лапласа

На практиці нас рідко цікавить ймовірність того, що дана подія відбудеться рівно m разів. Важливим буває оцінити ймовірність того, що ця кількість лежить у деяких межах. Таку оцінку можна отримати за допомогою наступної граничної теореми.

Теорема 3. (Інтегральна теорема Муавра-Лапласа.) *Нехай m - кількість появ події A в серії із n незалежних випробувань, p - ймовірність появи події A в окремому випробуванні, $0 < p < 1$; a і b будь-які фіксовані числа, $a < b$. Тоді*

$$\lim_{n\to\infty} P\left\{a \le \frac{m-np}{\sqrt{npq}} \le b\right\} = \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx\,, \tag{3.20}$$

причому наближення до граничного значення є рівномірним відносно a і b.

◁ Цю теорему можна довести за допомогою попередньої теореми, але ми будемо вважати її наслідком центральної граничної теореми (див. нижче).▷

Приклад 19. *Гральна кістка підкидається* 12000 *разів. Яка ймовірність того, що кількість випадань одиниці лежить між* 1900 *і* 2150 *?*

◁ Тут $n = 12000$, $p = 1/6$, $q = 5/6$, $\sqrt{npq} = 100/\sqrt{6}$, $m_1 = 1900$, $m_2 = 2150$. Тоді

$$\begin{aligned} &P\{m_1 \le m \le m_2\} = P\left\{\tfrac{m_1-np}{\sqrt{npq}} \le \tfrac{m-np}{\sqrt{npq}} \le \tfrac{m_2-np}{\sqrt{npq}}\right\} \approx \\ &\approx \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\sqrt{6}}^{\frac{3}{2}\sqrt{6}} e^{-\frac{x^2}{2}} dx = \tfrac{1}{\sqrt{2\pi}} \int\limits_0^{\sqrt{6}} e^{-\frac{x^2}{2}} dx + \tfrac{1}{\sqrt{2\pi}} \int\limits_0^{\frac{3}{2}\sqrt{6}} e^{-\frac{x^2}{2}} dx = \\ &= \Phi(\sqrt{6}) + \Phi(\tfrac{3}{2}\sqrt{6}) \approx 0,99\,. \end{aligned}$$

▷

Функція

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int\limits_0^z e^{-\frac{x^2}{2}} dx \tag{3.21}$$

називається *інтегралом похибок.*

Приклад 20. *Телефонна станція A обслуговує* 2000 *абонентів і повинна з'єднувати їх із іншою станцією B. Яка найменша кількість x ліній повинна зв'язувати A і B, щоб у* 99% *випадків викликів знайшлася вільна лінія? Нехай на протязі години пік кожний абонент розмовляє з B у середньому* 2 *хвилини.*

◁ Це є схема Бернуллі із $n = 2000$, $p = 1/30$ (ймовірність виклику). Число x визначається із умови: ймовірність того, що кількість викликів $\ge x$, повинна бути менше за $0,01$. Тобто

$$\begin{aligned} P\{m \ge x\} \;&\le\; 0,01 \Rightarrow P\{m\ge x\}{=}P\left\{\frac{m-np}{\sqrt{npq}} \ge \frac{x-np}{\sqrt{npq}}\right\} \approx \\ &\approx\; \frac{1}{2\pi} \int\limits_{\frac{x-np}{\sqrt{npq}}}^{\infty} e^{-\frac{x^2}{2}} dx = \frac{1}{2} - \Phi(\frac{x-np}{\sqrt{npq}}) \le 0,01\,. \end{aligned}$$

Звідси за таблицями знайдемо

$$\frac{x-np}{\sqrt{npq}} \ge 2.327 \Rightarrow x \ge 85.4\,.$$

Тобто $x = 86$. ▷

Розділ 4

Випадкові величини і функції розподілу

4.1. Випадкова величина

Для випадкової величини є характерним, що ми не можемо заздалегідь вказати значення, яке вона прийме, хоча, з іншого боку, множина її можливих значень вважається відомою. Ця множина може бути обмеженою або необмеженою. Проте для повного визначення випадкової величини необхідно вказати ще їх ймовірності, точніше ймовірності на множині значень.

Наприклад, для дискретних випадкових величин ξ, що приймають випадково те чи інше значення x (на дійсній прямій R_1) із зліченної множини X усіх можливих значень для ξ, можна визначити відповідні ймовірності

$$P_\xi(x) = P\{\xi = x\}, \quad x \in X\,. \tag{4.1}$$

Як приклад розглянемо схему незалежних випробувань і проаналізуємо випадкову величину - кількість успіхів. У цьому досліді простір елементарних подій Ω складається із 2^n елементарних подій ω - послідовностей типу $\omega = \{\omega_1\omega_2\omega_1\ldots\omega_1\}$. У схемі випробувань Бернуллі нас цікавили події A_k, $k = 0, 1, \ldots, n$, де A_k - ті послідовності, які містять у собі k успіхів ω_1. Таким чином, A_k містить C_n^k таких елементарних подій ω, і оскільки ймовірність кожного з них дорівнює $P(\{\omega\}) = p^k q^{n-k}$, то $P_n(A_k) = C_n^k p^k q^{n-k}$. Розглянемо функцію $\xi = \xi(\omega)$, що визначена на Ω за допомогою рівностей

$$\xi(\omega) = k, \quad \omega \in A_k, \quad k = 0, 1, \ldots, n\,. \tag{4.2}$$

Визначена таким чином функція $\xi(\omega)$ описує кількість успіхів у серії із n незалежних випробувань Бернуллі. Позначимо $\{\omega : \xi(\omega) = k\}$ множину тих ω, для яких $\xi(\omega) = k$. Тобто, за визначенням

$$A_k = \{\omega : \xi(\omega) = k\}, \quad P_n(\{\omega : \xi(\omega) = k\}) = C_n^k p^k q^{n-k}\,. \tag{4.3}$$

Звичайно використовується інший запис

$$P_n(k) = C_n^k p^k q^{n-k} = P_n\{\xi(\omega) = k\}, \quad k = 0, 1, \ldots, n\,. \tag{4.4}$$

Такий запис називається *розподілом випадкової величини* ξ. Таким чином, випадкова величина - кількість успіхів у серії із n незалежних випробувань Бернуллі має біноміальний розподіл.

У даному прикладі випадкова величина ξ приводить до множини $\overline{\Omega}$, яка складається з усіх значень величини ξ, тобто $\overline{\Omega} = \{0, 1, \ldots, n\}$. Алгебра усіх підмножин множини значень ξ складається з усіх підмножин множини $\overline{\Omega}$. Таким чином, із випадковою величиною ξ пов'язаний *новий ймовірнісний простір* $(\overline{\Omega}, \overline{F}, \overline{P})$, в якому простором елементарних подій $\overline{\Omega}$ є множина значень випадкової величини ξ, $\overline{F}$ - алгебра усіх підмножин $\overline{\Omega}$; ймовірність $\overline{P}$ зв'язана із ймовірністю P на початковому ймовірнісному просторі за допомогою формули (4.4): $\overline{P}(\{k\}) = P_n(k)$.

Випадкова величина $\xi = \xi(\omega)$ задає відображення простору (Ω, F, P) на простір $(\overline{\Omega}, \overline{F}, \overline{P})$. При цьому кожній точці $k \in \overline{\Omega}$ відповідає її прообраз в Ω - множині $\{\omega : \{\xi(\omega) = k\} \subset \Omega\}$. Задання $\xi(\omega)$ еквівалентно розбиттю Ω:

$$\Omega = \{\omega : \xi(\omega) = 0\} + \{\omega : \xi(\omega) = 1\} + \ldots + \{\omega : \xi(\omega) = n\}\,.$$

Твердження "ξ попадає в $\overline{A} \subset \overline{F}$" і "$\omega$ попадає в $A \subset F$" є еквівалентними. Є характерним те, що в теоретико-ймовірнісних задачах явна залежність $\xi = \xi(\omega)$ від ω, як правило, не відіграє ролі.

Розглянемо ще один приклад випадкової величини. Розглянемо простір елементарних подій Ω, який складається із нескінчених послідовностей $\omega = (\omega_1, \omega_2, \ldots)$ випробувань. Нехай ω містить k успіхів ω_1, а подія A_k складається із усіх таких ω (які k разів містять ω_1 - *процес Пуассона*). Покладемо за означенням

$$P(A_k) = \frac{\lambda^k e^{-\lambda}}{k!}\,, \quad k = 0, 1, \ldots \tag{4.5}$$

і визначимо випадкову величину ξ рівністю $\xi(\omega) = k$, $\omega \in A_k$. Ця випадкова величина має розподіл Пуассона, оскільки $P(\xi = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$, $\lambda > 0$, $k = 0, 1, \ldots$.

Це були приклади дискретних випадкових величин. У загальному випадку випадкова величина визначається таким чином.

Означення. Нехай (Ω, F, P) - ймовірнісний простір. *Випадковою величиною* ξ називається однозначна дійсна функція $\xi = \xi(\omega)$, що визначена на Ω, для якої множина елементарних подій вигляду $\{\omega : \xi(\omega) < x\}$ є подією (тобто належить F) для кожного дійсного числа $x \in R_1$. Таким чином, в означенні необхідно, щоб для кожного $x \in R_1$ множина $\{\omega : \xi = \xi(\omega)\} \in F$, і ця умова гарантує, що для кожного x визначена ймовірність події $\{\xi < x\} : F(x) = P\{\xi < x\}$. (Запис $\{\xi < x\}$ означає те ж саме, що і $\{\omega : \xi(\omega) < x\}$).

4.2. Функції розподілу

Функція $F(x) = P\{\xi < x\}\,, -\infty < x < \infty\,,$ називається *функцією розподілу* випадкової величини ξ.

Зазначимо, що функція $F(x)$ визначає ймовірність на множині значень ξ, але конструкція функції розподілу істотно простіша, ніж конструкція ймовірності.

Приклад 21. *Нехай ξ - кількість успіхів в серії з n незалежних випробувань Бернуллі. Тоді відповідна функція розподілу визначається рівністю*

$$F(x) = \begin{cases} 0 & x \le 0\,, \\ \sum\limits_{k<x} C_n^k p^k q^{n-k} & 0 < x \le n\,, \\ 1 & x > n\,. \end{cases} \tag{4.6}$$

Приклад 22. *Якщо ξ розподілена за законом Пуассона, то її функція розподілу має вигляд*

$$F(x) = \begin{cases} 0 & x \le 0\,, \\ \sum\limits_{k<x} \dfrac{\lambda^k e^{-\lambda}}{k!} & x > 0\,. \end{cases} \tag{4.7}$$

Означення. Випадкова величина ξ має *нормальний*, або *гауссовий, розподіл* $N(\mu, \sigma^2)$, якщо її функція розподілу має вигляд

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{x} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right] dz\,. \tag{4.8}$$

Розглянемо деякі властивості функції розподілу.
1. Для $x_2 > x_1$ маємо

$$\{\xi < x_2\} = \{\xi < x_1\} + \{x_1 \le \xi < x_2\}\,. \tag{4.9}$$

Оскільки події в правій частині (4.9) несумісні, то

$$P\{\xi < x_2\} = P\{\xi < x_1\} + P\{x_1 \le \xi < x_2\}\,, \tag{4.10}$$

або

$$P\{x_1 \le \xi < x_2\} = F(x_2) - F(x_1)\,. \tag{4.11}$$

2. Оскільки ліва частина (4.11) - невід'ємна, то звідси випливає, що $F(x)$ – неспадна функція для всіх $x \in R_1$. З означення $F(x)$ випливає, що

$$0 \le F(x) \le 1 \quad \text{для усіх } x \in R_1\,. \tag{4.12}$$

3. $F(x)$ неперервна зліва у кожній точці $x \in R_1$, тобто

$$F(x) = F(x-0) \equiv \lim_{x_k \uparrow x} F(x_k)\,. \tag{4.13}$$

Дійсно, нехай є деяка довільна послідовність $x_1 < x_2 < \ldots < x_n < \ldots$ та $\lim_{k\to\infty} x_k = x$. Подію $\xi < x$ можна представити у вигляді

$$\{\xi{<}x\} = \{\xi{<}x_1\} \bigcup \{\xi{<}x_2\} \bigcup \ldots = \{\xi{<}x_1\} + \{x_1{\le}\xi{<}x_2\} + \ldots$$

Внаслідок σ-адитивності ймовірності і рівності (4.11) маємо

$$\begin{aligned} F(x) &= F(x_1)+[F(x_2)-F(x_1)]+\ldots+[F(x_{k+1})-F(x_k)]+\ldots= \\ &= \lim_{k\to\infty} F(x_k)\,, \end{aligned}$$

і (4.13) доведено.

4. Праве граничне значення $F(x)$ в точці x дорівнює $P\{\xi \le x\}$, тобто

$$F(x+0) = \lim_{x_k \downarrow x} F(x_k) = P\{\xi \le x\}\,. \tag{4.14}$$

Дійсно, нехай є деяка довільна послідовність $x_1 > x_2 > \ldots > x_n > \ldots$ та $\lim_{k\to\infty} x_k = x$. Тоді на підставі (4.11) маємо

$$\begin{aligned} &\lim_{k\to\infty} F(x_k)=F(x_1)-[F(x_1)-F(x_2)]-\ldots-[F(x_k)-F(x_{k+1})]-\ldots= \\ &= P\{\xi<x_1\}-[P\{x_2\le\xi<x_1\}+\ldots+P\{x_{k+1}\le\xi<x_k\}+\ldots]= \\ &= P\{\xi < x_1\} - P\{x < \xi < x_1\} = P\{\xi \le x\}\,, \end{aligned}$$

і (4.14) доведено.

5. Справедливі такі співвідношення ($x_1 \le x_2$ - довільні):

$$P\{x_1 \le \xi \le x_2\} = F(x_2+0) - F(x_1)\,, \tag{4.15}$$

оскільки $\{\xi < x_1\} + \{x_1 \le \xi \le x_2\} = \{\xi \le x_2\}$. Зокрема, якщо $x_1 = x_2 = x$, то

$$P\{\xi = x\} = F(x+0) - F(x)\,. \tag{4.16}$$

А також мають місце співвідношення

$$P\{x_1 < \xi \le x_2\} = F(x_2+0) - F(x_1+0)\,, \tag{4.17}$$

оскільки $\{\xi \le x_1\} + \{x_1 < \xi \le x_2\} = \{\xi \le x_2\}$, і

$$P\{x_1 < \xi < x_2\} = F(x_2) - F(x_1+0)\,, \tag{4.18}$$

оскільки $\{\xi \le x_1\} + \{x_1 < \xi < x_2\} = \{\xi < x_2\}$.

6. Внаслідок неспадання $F(x)$ покладемо

$$F(-\infty) = \lim_{n\to\infty} F(-n), \quad F(\infty) = \lim_{n\to\infty} F(n)\,. \tag{4.19}$$

Тоді

$$F(-\infty) = 0\,, \quad F(\infty) = 1\,. \tag{4.20}$$

Дійсно

$$\begin{aligned} &F(+\infty) = \lim_{n\to\infty} F(n) = F(0) + [F(1) - F(0)] + \ldots + \\ &+[F(n+1) - F(n)] + \ldots = P\{\xi < 0\} + P\{0 \le \xi < 1\} + \ldots + \\ &+P\{n \le \xi < n+1\} + \ldots = P\{\xi < \infty\} = 1\,. \end{aligned}$$

Аналогічно
$$F(-\infty) = \lim_{n\to\infty} F(-n) =$$
$$= F(0) - [F(0) - F(-1)] - \ldots - [F(-n) - F(-n-1)] - \ldots =$$
$$= P\{\xi<0\} - [P\{-1\le\xi<0\} + \ldots + P\{-n-1\le\xi<-n\} + \ldots] = 0\,.$$

Можна показати, що $F(x)$ *неперервна* всюди на R_1 за винятком не більш ніж зліченної множини точок.

4.3. Дискретні випадкові величини

Означення. Випадкова величина ξ називається *дискретною*, якщо множина її значень скінченна або зліченна.

Для повної ймовірнісної характеристики дискретної випадкової величини, що приймає значення $x_1, x_2, \ldots$, достатньо задати ймовірності $p_k = P\{\xi = x_k\}$. Якщо значення x_k та p_k, $(k = 1, 2, \ldots)$ відомі, можна записати функцію розподілу $F(x)$ випадкової величини (дискретної) ξ у вигляді

$$F(x) = \sum_{k\,:\,x_k < x} p_k\,. \tag{4.21}$$

Очевидно, що $F(x)$ не залежить від способу нумерації значень випадкової величини ξ. Таким чином, функція розподілу випадкової величини зростає стрибками в точках $x = x_k$ і величина стрибка дорівнює

$$F(x_k + 0) - F(x_k) = p_k\,. \tag{4.22}$$

4.4. Неперервні випадкові величини

Означення. Випадкова величина ξ називається *неперервною*, якщо її функцію розподілу можна представити у вигляді

$$F(x) = \int_{-\infty}^{x} p(y)dy\,. \tag{4.23}$$

Функція $p(x), -\infty < x < \infty$ називається *густиною розподілу ймовірностей* (або *густиною ймовірностей*) випадкової величини ξ і припускається невід'ємною та кусочно-неперервною. (В літературі також використовується назва *щільність ймовірності*, але ми надалі будемо використовувати назву густина).

Густина $p(x)$ повністю визначає функцію розподілу і в точках неперервності має місце співвідношення

$$p(x) = \frac{dF(x)}{dx}\,. \tag{4.24}$$

Для будь-яких $x_1 < x_2$ має місце

$$P\{x_1 \le \xi < x_2\} = F(x_2) - F(x_1) = \int\limits_{x_1}^{x_2} p(y)dy\,. \tag{4.25}$$

Зокрема

$$P\{x \le \xi < x + \delta x\} = p(x) \cdot \delta x + O\left((\delta x)^2\right)\,, \tag{4.26}$$

у тому числі для неперервної випадкової величини

$$P\{\xi = x\} = F(x+0) - F(x) = 0\,. \tag{4.27}$$

Очевидно, що

$$\int\limits_{-\infty}^{\infty} p(x)dx = F(+\infty) = 1\,. \tag{4.28}$$

Прикладом неперервної випадкової величини є нормальна $N(\mu, \sigma^2)$ випадкова величина з густиною

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}\,. \tag{4.29}$$

Все, що сказано вище про функцію розподілу, автоматично переноситься на випадок умовної ймовірності. Якщо $P(B) > 0$, то величина $F(x|B) = P\{\xi < x|B\}$ називається *умовною функцією розподілу* випадкової величини ξ. Вона має усі вказані вище властивості функції розподілу.

4.5. Багатовимірні (векторні) випадкові величини

У задачах із випадковим елементарним результатом часто доводиться враховувати взаємодію різних випадкових факторів.

Означення. Нехай (Ω, F, P) - ймовірнісний простір, і нехай $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$ - випадкові величини, що визначені на Ω. Вектор $\vec{\xi}(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega))$ називається *випадковим вектором* або *n-вимірною випадковою величиною*, а $\xi_i(\omega)$, $i = 1, \ldots, n$ називаються *координатами* або *компонентами* випадкового вектора $\vec{\xi}$.

Оскільки всі $\xi_i(\omega)$ задані на одному і тому ж ймовірнісному просторі Ω, а F замкнуто відносно добутків будь-якої скінченої кількості подій, то множина $\{\omega : \xi_1(\omega) < x_1, \ldots, \xi_n(\omega) < x_n\} \in F$ для будь-яких $x_1, \ldots, x_n \in R_1$.

Означення. Функція $F(x_1, \ldots, x_n) = P\{\xi_1 < x_1, \ldots, \xi_n < x_n\}$ називається *n-вимірною функцією розподілу* випадкової величини $\vec{\xi} = (\xi_1, \ldots, \xi_n)$.

Заради наочності нижче будемо в основному розглядати двовимірні випадкові величини. Функція

$$F(x,y) = P\{\xi < x, \eta < y\} \tag{4.30}$$

задає ймовірність того, що точка з координатами (ξ, η) попаде у нескінчений прямокутник (Рис.4.1А).

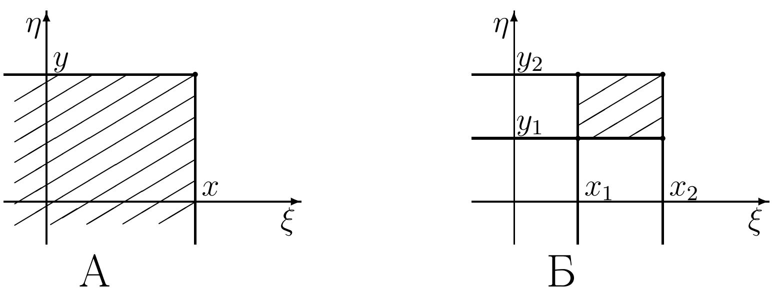

Рис. 4.1.

Основні властивості двовимірної функції розподілу:
1) $F(x,y)$ є неспадною по x і y.
2) $F(x,y)$ неперервна зліва по x і y.

3) $$F(\infty,\infty) = 1,\ F(-\infty,y) = 0,\ F(x,-\infty) = 0\,. \tag{4.31}$$

Доведення цих властивостей аналогічно наведеним вище для одновимірної функції розподілу.

4) $$\begin{aligned} P\{x_1 \le \xi < x_2, y_1 \le \eta < y_2\} &= F(x_2,y_2) - F(x_1,y_2) - \\ &- F(x_2,y_1) + F(x_1,y_1)\,. \end{aligned} \tag{4.32}$$

Доведення випливає із (4.30) (Див. Рис. 4.1Б).
5) Можна отримати розподіли координат ξ та η (т.зв. *маргінальні розподіли*):

$$F_\xi(x) = F(x,\infty),\ F_\eta(y) = F(\infty,y)\,. \tag{4.33}$$

$\lhd$ Очевидно, що можна записати для подій

$$\{\xi < x\} = \sum_{k=-\infty}^{\infty} \{\xi < x, k \le \eta < k+1\}\,. \tag{4.34}$$

Внаслідок попарної несумісності подій у правій частині (4.34) маємо

$$\begin{aligned}
F_\xi(x) &= P\{\xi < x\} = \sum_{k=-\infty}^{\infty} P\{\xi < x, k \le \eta < k+1\} = \\
&= \sum_{k=-\infty}^{\infty} [F(x,k+1) - F(x,k)] = \\
&= \lim_{N_1,N_2\to\infty} \sum_{k=-N_1}^{N_2} [F(x,k+1) - F(x,k)] = \\
&= \lim_{N_1,N_2\to\infty} [F(x,N_2) - F(x,-N_1)] = \\
&= F(x,\infty) - F(x,-\infty) = F(x,\infty)\,.
\end{aligned}$$

$\triangleright$

Означення. Випадковий вектор називається *дискретним*, якщо кожна його координата - дискретна випадкова величина, і *неперервним*, якщо існує кусочно-неперервна невід'ємна функція $p(x,y)$, $x, y \in R_1$ така, що для будь-яких x і y є справедливою рівність:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(z_1,z_2)dz_1dz_2\,. \tag{4.35}$$

Функція $p(x,y)$ називається *густиною ймовірності* випадкового вектора (ξ,η).

Густина ймовірності випадкового вектора має такі властивості:

1) У точках неперервності $p(x,y)$ справедлива рівність

$$p(x,y) = \frac{\partial^2 F}{\partial x \partial y}\,. \tag{4.36}$$

Оскільки $F(x,y)$ не спадає по кожній змінній, то $p(xy) \ge 0$.

2) Для будь-якої вимірної області $D \subset R_2$ маємо

$$P\{(\xi,\eta) \in D\} = \iint_D p(x,y)dxdy\,. \tag{4.37}$$

Зокрема, за допомогою теореми про середнє,

$$P\{x \le \xi < x+dx,\ y \le \eta < y+dy\} = p(x,y)dxdy + o(dxdy)\,.$$

3) Внаслідок $F(+\infty,+\infty) = 1$ маємо

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)dxdy = 1\,. \tag{4.38}$$

4) На основі (4.33) і (4.35) робимо висновок, що якщо двовимірна випадкова величина (ξ,η) має густину, то кожна її компонента має густину, причому

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)dy, \quad p_\eta(y) = \int_{-\infty}^{\infty} p(x,y)dx\,. \tag{4.39}$$

Приклад 23. *Випадковий вектор (ξ, η) називається* **рівномірно розподіленим** *в області D, якщо*

$$p(x,y) = \begin{cases} \dfrac{1}{\mu(D)}, & (x,y) \in D\,, \\ 0\,, & (x,y) \notin D\,, \end{cases} \tag{4.40}$$

де $\mu(D)$ - площа області D. Нехай D^* - вимірна область на площині. Тоді

$$P\{(\xi,\eta) \in D^*\} = \frac{\mu(D^* \bigcap D)}{\mu(D)} \tag{4.41}$$

(Це окремий випадок геометричного розподілу).

Приклад 24. *Випадковий вектор $(\xi_1, \ldots, \xi_n)$ називається* **нормально розподіленим**, *якщо його густина дорівнює*

$$p(x_1,\ldots,x_n) = \left\{\frac{detA}{(2\pi)^n}\right\}^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum_{i,j} a_{ij}(x_i-\mu_i)(x_j-\mu_j)\right\}, \tag{4.42}$$

де $A = \|a_{ij}\|$ додатньо визначена матриця (обернена до **матриці коваріацій** *(див. нижче)). Зокрема,* **коефіцієнт кореляції** $r = \dfrac{cov(\xi,\eta)}{\sigma_1\sigma_2}$.

Нехай густина $p(x,y)$ випадкового вектора (ξ,η) неперервна. Розглянемо подію $B = \{y \le \eta < y + \Delta y\}$, тоді

$$P(B) = P\{y \le \eta < y + \Delta y\} = \int\limits_y^{y+\Delta y} p_\eta(z)dz\,,$$

де $p_\eta(z) = \int_{-\infty}^{\infty} p(x,z)dx$. Аналогічно можна записати

$$P\{\xi < x, y \le \eta < y + \Delta y\} = \int\limits_{-\infty}^{x} \int\limits_{y}^{y+\Delta y} p(t,z)dtdz\,.$$

Тоді, внаслідок визначення умовної ймовірності, якщо $P(B)>0$, то

$$F_\xi(x|B) = P\{\xi<x|B\} = \frac{P\{\xi<x, y\le\eta<y+\Delta y\}}{P(B)} = \frac{\int\limits_{-\infty}^{x}\int\limits_{y}^{y+\Delta y} p(t,z)dtdz}{\int\limits_{y}^{y+\Delta y} p_\eta(z)dz}\,. \tag{4.43}$$

Продиференціюємо (4.43) по x і спрямуємо $\Delta y \to 0$:

$$p_\xi(x|y) = \frac{dF_\xi(x|\eta = y)}{dx} = \frac{p(x,y)}{p_\eta(y)}\,, \quad p_\eta(y) \neq 0\,. \tag{4.44}$$

Фунція $p_\xi(x|y)$ називається *густиною ймовірності умовного розподілу* ξ при умові $\eta = y$. Із (4.44) випливає

$$p(x,y) = p_\xi(x|y)p_\eta(y)\,, \tag{4.45}$$

що за формою нагадує теорему множення ймовірностей. Спрямуємо в (4.43) $\Delta y \to 0$ і скористуємось визначенням (4.44)

$$F_\xi(x|\eta = y) = P\{\xi < x|\eta = y\} = \int\limits_{-\infty}^{x} p_\xi(t|y)dt\,. \tag{4.46}$$

Проінтегруємо (4.45) по y і використаємо визначення (4.39)

$$p_\xi(x) = \int\limits_{-\infty}^{\infty} p_\eta(y)p_\xi(x|y)dy\,. \tag{4.47}$$

Ця рівність є неперервним аналогом формули повної ймовірності.

4.6. Незалежність випадкових величин

Означення. Випадкові величини $\xi_1, \ldots, \xi_n$ називаються *незалежними* (у сукупності), якщо для довільних $x_1, \ldots, x_n$ події $\{\xi_1 < x_1\}, \ldots, \{\xi_n < x_n\}$ незалежні у сукупності, тобто

$$P\left(\{\xi_1 < x_1\}\bigcap \ldots \bigcap\{\xi_n < x_n\}\right) = P\{\xi_1 < x_1\}\cdots P\{\xi_n < x_n\}\,,$$

або

$$F(x_1, \ldots, x_n) = F_{\xi_1}(x_1)\cdots F_{\xi_n}(x_n)\,. \tag{4.48}$$

Для незалежних випадкових величин ξ і η маємо для довільних $x_1 < x_2$, $y_1 < y_2$

$$\begin{aligned}
&P\{x_1 \le \xi < x_2, y_1 \le \eta < y_2\} = \\
&= P\{\xi{<}x_2, \eta{<}y_2\} - P\{\xi{<}x_1, \eta{<}y_2\} - P\{\xi{<}x_2, \eta{<}y_1\} + \\
&+ P\{\xi{<}x_1, \eta{<}y_1\} = P\{\xi{<}x_2\}\left[P\{\eta{<}y_2\} - P\{\eta{<}y_1\}\right] - \\
&- P\{\xi{<}x_1\}\left[P\{\eta{<}y_2\} - P\{\eta{<}y_1\}\right] = P\{x_1{\le}\xi{<}x_2\}P\{y_1{\le}\eta{<}y_2\}\,.
\end{aligned}$$

Якщо незалежні випадкові величини ξ та η мають відповідні густини ймовірностей $p_\xi(x)$ та $p_\eta(y)$, то вектор (ξ, η) має густину

$$p(x,y) = p_\xi(x)p_\eta(y)\,. \tag{4.49}$$

Ця рівність є наслідком (4.48) у точках (x,y), де $p_\xi(x)$ та $p_\eta(y)$ неперервні, в інших точках вона є визначенням.

Важливим є зворотнє *твердження*. Якщо виконується (4.49), то ξ та η незалежні, оскільки виконується (4.48).

Приклад 25. *Нехай $\xi \in N(\mu_1, \sigma_1^2)$, $\eta \in N(\mu_2, \sigma_2^2)$ і незалежні. Тоді густина випадкового вектора відповідно до (4.49) дорівнює добутку маргінальних густин:*

$$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}. \tag{4.50}$$

Вираз (4.50) - це окремий випадок (4.42), коли

$$A = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix}, \quad \det A = \frac{1}{\sigma_1^2\sigma_2^2}, \quad r = 0\,. \tag{4.51}$$

4.7. Функції від випадкових величин

Обмежимось прикладом двох випадкових величин. Нехай на ймовірнісному просторі (Ω, F, P) визначені дві випадкові величини $\xi_1 = \xi_1(\omega)$ та $\xi_2 = \xi_2(\omega)$, $\omega \in \Omega$. Нехай $F(x_1, x_2)$ - функція розподілу випадкового вектора (ξ_1, ξ_2). Розглянемо деякі функції η_1 та η_2, тобто нові випадкові величини

$$\begin{aligned} \eta_1 &= f_1(\xi_1,\xi_2) = f_1\left(\xi_1(\omega), \xi_2(\omega)\right), \\ \eta_2 &= f_2(\xi_1,\xi_2) = f_2\left(\xi_1(\omega), \xi_2(\omega)\right). \end{aligned} \tag{4.52}$$

Таким чином, η_1 та η_2 є складними функціями ω, що задані на Ω. Необхідно за відомою функцією розподілу $F(x_1, x_2)$ знайти функцію розподілу $\Phi(y_1, y_2)$ випадкового вектора (η_1, η_2). Маємо

$$\begin{aligned} \Phi(y_1,y_2) &= P\{\eta_1 < y_1, \eta_2 < y_2\} = \\ &= P\left\{f_1(\xi_1,\xi_2) < y_1, f_2(\xi_1,\xi_2) < y_2\right\}. \end{aligned} \tag{4.53}$$

1) Розглянемо дискретні випадкові величини. Нехай

$$p_{ij} = P\{\xi_1 = x_{1i}, \xi_2 = x_{2j}\}, \quad i,j = 1,2,\ldots \tag{4.54}$$

Тоді, відповідно до (4.53), маємо

$$\Phi(y_1,y_2) = \sum_{i,j\in\mathcal{U}} p_{ij}\,, \tag{4.55}$$

де множина індексів $\mathcal{U}$ така, що

$$\mathcal{U} = \{(i,j) : f_1(x_{1i}, x_{2j}) < y_1, f_2(x_{1i}, x_{2j}) < y_2\}\,. \tag{4.56}$$

2) Розглянемо неперервні випадкові величини. Нехай $p(x_1, x_2)$ - густина ймовірності вектора (ξ_1, ξ_2). Тоді, відповідно до (4.53) та (4.37), маємо

$$\Phi(y_1,y_2) = \iint_D p(x_1,x_2)dx_1dx_2\,, \tag{4.57}$$

де область D така, що

$$D = \{(x_1, x_2) : f_1(x_1, x_2) < y_1, f_2(x_1, x_2) < y_2\} \,. \tag{4.58}$$

Приклад 26. *Знайти функцію розподілу суми $\eta = \xi_1 + \xi_2$, якщо задана густина розподілу ймовірності $p(x_1, x_2)$ випадкового вектора (ξ_1, ξ_2).*

$\lhd$ Маємо

$$\begin{aligned}\Phi(y) &= P\{\eta{<}y\}{=}P\{\xi_1{+}\xi_2{<}y\} = \iint\limits_{x_1+x_2<y} p(x_1, x_2)dx_1dx_2 = \\ &= \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{y-x_1} dx_2 p(x_1, x_2) = \int\limits_{-\infty}^{\infty} dx_2 \int\limits_{-\infty}^{y-x_2} dx_1 p(x_1, x_2)\,.\end{aligned}$$

Зробимо заміну змінних $z = x_1 + x_2$, $x_1 = x_1$ (якобіан дорівнює одиниці) і отримаємо

$$\Phi(y) = \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{y} dz p(x_1, z - x_1)\,. \tag{4.59}$$

Із (4.59) випливає, що густина ймовірності випадкової величини η дорівнює

$$p_\eta(y) = \Phi'(y) = \int\limits_{-\infty}^{\infty} dx\, p(x, y - x)\,. \tag{4.60}$$

Якщо при цьому випадкові величини ξ_1 та ξ_2 незалежні, то $p(x_1, x_2) = p_{\xi_1}(x_1)p_{\xi_2}(x_2)$ та

$$p_\eta(y){=}\int\limits_{-\infty}^{\infty} dx\, p_{\xi_1}(x)p_{\xi_2}(y{-}x){=}\int\limits_{-\infty}^{\infty} dx\, p_{\xi_1}(y{-}x)p_{\xi_2}(x){=}p_{\xi_1} * p_{\xi_2}\,. \tag{4.61}$$

Визначення законів розподілу суми по законам розподілу незалежних доданків називається *композицією (згорткою)* законів розподілу доданків. $\rhd$

Теорема 4. *Нехай ξ_1 та ξ_2 незалежні випадкові величини, і $f_1(x)$ та $f_2(x)$ довільні функції, такі, що $\eta_1 = f_1(\xi_1)$ та $\eta_2 = f_2(\xi_2)$ також випадкові величини. Тоді η_1 та η_2 незалежні, тобто функції від незалежних випадкових величин є незалежними.*

$\lhd$ Розглянемо дискретні випадкові величини. Нехай $\xi_1 = x_1, x_2, \ldots$ та $\xi_2 = y_1, y_2, \ldots$. Тоді випадкові величини $\eta_1 = f_1(\xi_1)$ та $\eta_2 = f_2(\xi_2)$ також дискретні. Нехай λ_1 та λ_1 довільні фіксовані значення η_1 та η_2. Тоді на підставі (4.49) маємо

$$\begin{aligned}&P\{f_1(\xi_1) = \lambda_1, f_2(\xi_2) = \lambda_2\} = \sum_{k,l} P\{\xi_1 = x_k, \xi_2 = y_l\} = \\ &= \sum_{k,l} P\{\xi_1 = x_k\}P\{\xi_2 = y_l\} = \sum_{k} P\{\xi_1 = x_k\} \sum_{l} P\{\xi_2 = y_l\} = \\ &= P\{f_1(\xi_1) = \lambda_1\}\, P\{f_2(\xi_2) = \lambda_2\}\,,\end{aligned}$$

де підсумовування відбувається по тих значеннях k та l, що $f_1(x_k) = \lambda_1$ та $f_2(y_l) = \lambda_2$. Звідси випливає незалежність випадкових величин $\eta_1 = f_1(\xi_1)$ та $\eta_2 = f_2(\xi_2)$. $\rhd$

Приклад 27. *Густини розподілів для лінійних перетворень.*

$\lhd$ Нехай $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ та $\vec{\eta} = (\eta_1, \ldots, \eta_n)$ - n-вимірні випадкові величини, причому $\vec{\eta} = A\vec{\xi}$, тобто $\eta_i = \sum_j a_{ij}\xi_j$, де A - невироджена матриця. Нехай $p_{\vec{\xi}}(\vec{x}) = p_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$ - густина розподілу випадкового вектора $\vec{\xi}$, D - довільна квадрована область в R_n. Тоді

$$\begin{aligned}\int\limits_D p_{\vec{\eta}}(\vec{y})d\vec{y} &= P\{\vec{\eta} \in D\} = P\{A\vec{\xi} \in D\} = P\{\vec{\xi} \in A^{-1}D\} = \\ &= \int\limits_{A^{-1}D} p_{\vec{\xi}}(\vec{x})d\vec{x} = \int\limits_D p_{\vec{\xi}}(A^{-1}\vec{y})|\det A^{-1}|d\vec{y}.\end{aligned}$$

Внаслідок довільності області D маємо зв'язок густин випадкових векторів $\vec{\xi}$ та $\vec{\eta} = A\vec{\xi}$

$$p_{\vec{\eta}}(\vec{x}) = p_{\vec{\xi}}(A^{-1}\vec{x})|\det A^{-1}|\,. \tag{4.62}$$

$\rhd$

Приклад 28. *Нехай (ξ, η) - неперервний випадковий вектор із густиною ймовірностей $p(x, y)$. Знайти функцію розподілу добутку $\chi = \xi\eta$.*

$\lhd$ Маємо

$$\begin{aligned}F_\chi(z) &= P\left\{\xi\eta < z\right\} = \iint\limits_{xy < z} p(x, y)dxdy = \\ &= \int\limits_{-\infty}^{0} dx \int\limits_{z/x}^{\infty} dy\, p(x, y) + \int\limits_{0}^{\infty} dx \int\limits_{-\infty}^{z/x} dy\, p(x, y)\,.\end{aligned}$$

Тоді

$$p_\chi(z) = F'_\chi(z) = -\int\limits_{-\infty}^{0} dx \cdot \frac{1}{x} \cdot p(x, \frac{z}{x}) + \int\limits_{0}^{\infty} dx \cdot \frac{1}{x} \cdot p(x, \frac{z}{x})\,. \tag{4.63}$$

$\rhd$

Приклад 29. *Нехай (ξ, η) - неперервний випадковий вектор із густиною ймовірностей $p(x, y)$. Знайти функцію розподілу відношення $\chi = \xi/\eta$.*

$\lhd$ Маємо

$$\begin{aligned}F_\chi(z) &= P\left\{\frac{\xi}{\eta} < z\right\} = \iint\limits_{\frac{x}{y} < z} p(x, y)dxdy = \\ &= \int\limits_{-\infty}^{0} dy \int\limits_{zy}^{\infty} dx\, p(x, y) + \int\limits_{0}^{\infty} dy \int\limits_{-\infty}^{zy} dx\, p(x, y)\,.\end{aligned}$$

Тоді

$$p_\chi(z) = F'_\chi(z) = -\int\limits_{-\infty}^{0} dy \cdot y \cdot p(zy, y) + \int\limits_{0}^{\infty} dy \cdot y \cdot p(zy, y)\,. \tag{4.64}$$

Нехай ξ та η незалежні та розподілені нормально, тобто $\xi, \eta \in N(0, 1)$. У цьому випадку

$$p_\chi(z) = -\frac{1}{2\pi}\int\limits_{-\infty}^{0} dy \cdot y \cdot e^{-\frac{1}{2}(1+z^2)y^2} + \frac{1}{2\pi}\int\limits_{0}^{\infty} dy \cdot y \cdot e^{-\frac{1}{2}(1+z^2)y^2} =$$
$$= \frac{1}{\pi}\int\limits_{0}^{\infty} dy \cdot y \cdot e^{-\frac{1}{2}(1+z^2)y^2} = \frac{1}{\pi} \cdot \frac{1}{1+z^2}\,.$$

Ми отримали т.зв. *розподіл Коші* (у фізиці він часто називається *розподілом Лорентца*). $\triangleright$

Розділ 5

Числові характеристики випадкових величин

5.1. Моменти випадкових величин

Означення. Моментом порядку k дискретної випадкової величини ξ, що приймає значення x_i з ймовірністю $p\{\xi = x_i\} = p_i$, $i = 1, 2, \ldots$, називається число

$$M\xi^k = \sum_{i=1}^{\infty} x_i^k p_i \tag{5.1}$$

при умові, що ряд (5.1) збігається абсолютно, тобто

$$M|\xi^k| = \sum_{i=1}^{\infty} |x_i^k| p_i < \infty\,. \tag{5.2}$$

Величина $M|\xi^k|$ називається *абсолютним моментом порядку k.*

Моментом порядку k неперервної випадкової величини ξ із густиною $p(x)$ називається число

$$M\xi^k = \int\limits_{-\infty}^{\infty} x^k p(x) dx \tag{5.3}$$

при умові, що інтеграл збігається абсолютно, тобто

$$M|\xi^k| = \int\limits_{-\infty}^{\infty} |x^k| p(x) dx < \infty\,. \tag{5.4}$$

Окремий випадок. При $k = 1$ (5.1) та (5.3) визначають *математичне сподівання* (*середнє значення*) випадкової величини ξ.

Приклад 30. *Розподіл Пуассона.*

$\triangleleft$ Випадкова величина ξ приймає значення $\xi = k$ з ймовірністю

$$p_k = P\{\xi = k\} = \frac{\lambda^k}{k!}e^{-\lambda},\ \lambda > 0,\ k = 0, 1, 2, \dots$$

Тоді

$$M\xi = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = \lambda\,. \tag{5.5}$$

$\triangleright$

Приклад 31. *Рівномірний розподіл.*

$\triangleleft$ Випадкова величина ξ рівномірно розподілена на інтервалі $[a, b]$, тобто

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

Тоді

$$M\xi = \frac{1}{b-a}\int_a^b x\,dx = \frac{a+b}{2} \quad \text{- середина відрізку } [a, b]\,.$$

$\triangleright$

Приклад 32. *Розподіл Коші.*

$\triangleleft$ Тут $p(x) = \dfrac{1}{\pi(1+x^2)}$ та $M|\xi| = \frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{|x|}{1+x^2}dx = \infty$. Тобто $M\xi$ не існує. $\triangleright$

Приклад 33. *Нормальний розподіл.*

$\triangleleft$ Тут $p(x) = \dfrac{1}{\sqrt{2\pi}\sigma}\exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$. Тоді

$$M\xi = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx = \mu\,. \tag{5.6}$$

$\triangleright$

Теорема 5. *Якщо ξ - дискретна (неперервна) випадкова величина, що приймає значення $x_1, x_2, \dots$ з ймовірностями $p_1, p_2, \dots$, а $\eta = f(\xi)$ нова випадкова величина, то*

$$M\eta = Mf(\xi) = \sum_{i=1}^{\infty} f(x_i)p_i \left(= \int_{-\infty}^{\infty} f(x)p(x)dx \right)\,, \tag{5.7}$$

якщо ряд (інтеграл) (5.7) збігається абсолютно.

◁ Для дискретної випадкової величини ξ маємо дискретну випадкову величину $\eta = f(\xi)$, що приймає значення $y_1 = f(\xi_1)$, $y_2 = f(\xi_2), \dots$ з ймовірностями q_1, $q_2, \dots$

$$q_s = \overline{P}\{\eta = y_s\} = P\left\{\omega : f\left(\xi(\omega)\right) = y_s\right\} = \sum_{k: f(x_k) = y_s} p_k \,.$$

Тоді маємо

$$\begin{aligned} M\eta &= \sum_{s=1}^{\infty} y_s q_s = \sum_{s=1}^{\infty} y_s \sum_{k: f(x_k)=y_s} p_k = \\ &= \sum_{s=1}^{\infty} \left\{ \sum_{k: f(x_k)=y_s} f(x_k) p_k \right\} = \sum_{i=1}^{\infty} f(x_i) p_i \,. \end{aligned} \tag{5.8}$$

Співвідношення (5.8) виконується внаслідок того, що кожний доданок $f(x_i)p_i$ приймає участь у двох останніх сумах тільки один раз, оскільки всі y_s різні. Можливість об'єднання в одну суму випливає із леми про підсумовування по блоках, оскільки ряд $\sum_i |f(x_i)| p_i$ збігається за умовою. Аналогічне доведення для неперервної випадкової величини. ▷

Означення. *k-м центральним моментом* називається математичне сподівання

$$M(\xi - M\xi)^k \,, \tag{5.9}$$

якщо існує $M|\xi - M\xi|^k$.

Означення. Центральний момент другого порядку називається *дисперсією*

$$D\xi = M(\xi - M\xi)^2 \,. \tag{5.10}$$

Очевидно, що $M(\xi - M\xi)^2 = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2$, тобто

$$D\xi = M\xi^2 - (M\xi)^2 \,. \tag{5.11}$$

Означення. Величина

$$\sigma\xi = \sqrt{D\xi} \tag{5.12}$$

називається *середньоквадратичним відхиленням* ξ.

На підставі доведенної вище теореми і означення (5.10) можна записати

$$D\xi = \sum_{k=1}^{\infty} (x_k - M\xi)^2 p_k \,, \tag{5.13}$$

$$D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 p(x) dx \,. \tag{5.14}$$

Приклад 34. *Нормальний розподіл $N(\mu, \sigma^2)$.*

$\triangleleft$ Оскільки, відповідно до (5.6) $M\xi = \mu$, то

$$D\xi = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\infty} (x-\mu)^2 \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = \sigma^2 \,. \tag{5.15}$$

Нормальний розподіл повністю визначається параметрами μ та σ. $\triangleright$

Приклад 35. *Розподіл Пуассона:* $P\{\xi = k\} = \dfrac{\lambda^k}{k!} e^{-\lambda}$, $\lambda > 0$.

$\triangleleft$ Оскільки, відповідно до (5.5), $M\xi = \lambda$, то

$$\begin{aligned} D\xi &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 = e^{-\lambda} \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} + e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} - \lambda^2 = \\ &= \lambda^2 e^{-\lambda} \frac{d^2}{d\lambda^2} (\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}) + \lambda e^{-\lambda} \frac{d}{d\lambda} (\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}) - \lambda^2 = \\ &= e^{-\lambda} (\lambda^2 e^{\lambda} + \lambda e^{\lambda}) - \lambda^2 = \lambda \,. \end{aligned}$$

Таким чином, єдиний параметр λ визначає математичне сподівання і дисперсію розподілу Пуассона. $\triangleright$

5.2. Властивості математичного сподівання і дисперсії

1) За умови, що $M|\xi|$ та $M|\eta|$ скінченні, маємо

$$M(\xi + \eta) = M\xi + M\eta \,. \tag{5.16}$$

$\triangleleft$ Наведемо доведення тільки для неперервної випадкової величини (ξ, η) з густиною $p(x, y)$. Відповідно до (4.60) маємо для $\chi = \xi + \eta$ густину $p_\chi(z) = \int\limits_{-\infty}^{\infty} dx\, p(x, z - x)$. Тому

$$\begin{aligned} M\chi &= \int\limits_{-\infty}^{\infty} z p_\chi(z) dz = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} z p(x, z - x) dx dz = \\ &= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x + y) p(x, y) dx dy = \\ &= \int\limits_{-\infty}^{\infty} dx\, x \int\limits_{-\infty}^{\infty} dy\, p(x, y) + \int\limits_{-\infty}^{\infty} dy\, y \int\limits_{-\infty}^{\infty} dx\, p(x, y) = \\ &= \int\limits_{-\infty}^{\infty} dx\, x\, p_\xi(x) + \int\limits_{-\infty}^{\infty} dy\, y\, p_\eta(y) = M\xi + M\eta \,. \end{aligned}$$

Порядок інтегралів можна змінювати внаслідок їх абсолютної збіжності. $\triangleright$
2) Із властивості 1) випливає

$$M(c_1\xi + c_2\eta) = c_1 M\xi + c_2 M\eta\,. \tag{5.17}$$

3) Якщо випадкові величини ξ та η незалежні, то

$$M(\xi \cdot \eta) = M\xi \cdot M\eta \tag{5.18}$$

за умови, що $M|\xi|$ та $M|\eta|$ скінченні.
$\triangleleft$ Наведемо доведення тільки для неперервної випадкової величини (ξ, η) із густиною $p(x, y)$. Відповідно до (4.63) маємо для $\chi = \xi\eta$

$$p_\chi(z) = -\int\limits_{-\infty}^{0} dx \cdot \frac{1}{x} \cdot p_\xi(x) p_\eta\Big(\frac{z}{x}\Big) + \int\limits_{0}^{\infty} dx \cdot \frac{1}{x} \cdot p_\xi(x) p_\eta\Big(\frac{z}{x}\Big)\,. \tag{5.19}$$

Звідси

$$\begin{aligned} M\chi &= \int\limits_{-\infty}^{\infty} z \cdot p_\chi(z) \cdot dz = \\ &= -\int\limits_{-\infty}^{0} dx \cdot \frac{1}{x} \cdot p_\xi(x) \int\limits_{-\infty}^{\infty} dz \cdot p_\eta\Big(\frac{z}{x}\Big) \cdot z + \int\limits_{0}^{\infty} dx \cdot \frac{1}{x} \cdot p_\xi(x) \int\limits_{-\infty}^{\infty} dz \cdot p_\eta\Big(\frac{z}{x}\Big) \cdot z = \\ &= \int\limits_{-\infty}^{0} dx \cdot \frac{1}{x} \cdot p_\xi(x) \int\limits_{-\infty}^{\infty} dt \cdot p_\eta(t) t x^2 + \int\limits_{0}^{\infty} dx \cdot \frac{1}{x} \cdot p_\xi(x) \int\limits_{-\infty}^{\infty} dt \cdot p_\eta(t) t x^2 = \\ &= \int\limits_{-\infty}^{\infty} dx \cdot p_\xi(x) x \cdot \int\limits_{-\infty}^{\infty} dt \cdot p_\eta(t) t = M\xi \cdot M\eta\,. \end{aligned}$$

Порядок інтегралів можна змінювати внаслідок їх абсолютної збіжності. $\triangleright$
4) Якщо $\xi \ge \eta$, то

$$M(\xi) \ge M\eta\,. \tag{5.20}$$

5) *Нерівність Коші-Буняковського*:

$$M|\xi \cdot \eta| \le \sqrt{M\xi^2 \cdot M\eta^2}\,, \tag{5.21}$$

якщо величини праворуч скінченні.
$\triangleleft$ Маємо $|\xi\eta| \le \frac{1}{2}(\xi^2 + \eta^2)$. Тоді, якщо скінченні $M\xi^2$ та $M\eta^2$, то внаслідок (5.20) скінченно $M|\xi\eta|$. При довільному λ маємо

$$0 \le M(\lambda|\xi| + |\eta|)^2 = \lambda^2 M\xi^2 + 2\lambda M|\xi\eta| + M\eta^2\,. \tag{5.22}$$

Оскільки це справедливо при будь-якому λ, то дискримінант повинен бути недодатнім, тобто

$$(M|\xi \cdot \eta|)^2 - M\xi^2 \cdot M\eta^2 \le 0\,.$$

$\triangleright$

6) *Нерівність Чебишева*: Для будь-якого $\varepsilon > 0$ є справедливою оцінка

$$P\{|\xi| > \varepsilon\} \le \frac{M|\xi|^2}{\varepsilon^2}\,, \tag{5.23}$$

якщо величина $M|\xi|^2$ є скінченною.
$\triangleleft$ Визначимо випадкову величину η за допомогою співвідношень

$$\eta = \begin{cases} 0, & \text{якщо } |\xi| \le \varepsilon\,, \\ \varepsilon, & \text{якщо } |\xi| > \varepsilon\,. \end{cases} \tag{5.24}$$

Таким чином, η - дискретна випадкова величина, що приймає два значення: 0 з ймовірністю $p = P\{|\xi| \le \varepsilon\}$ та ε з ймовірністю $q = 1 - p = P\{|\xi| > \varepsilon\}$. Із визначення η випливає, що $\eta^2 \le |\xi|^2$, тоді внаслідок (5.20) маємо $M|\xi|^2 \ge M\eta^2 = \varepsilon^2 P\{|\xi| > \varepsilon\}$, звідки випливає (5.23). $\triangleright$
Зробимо в (5.23) заміну $\xi \to \xi - M\xi$, тоді

$$P\{|\xi - M\xi| > \varepsilon\} \le \frac{M(\xi - M\xi)^2}{\varepsilon^2} = \frac{D\xi}{\varepsilon^2}\,. \tag{5.25}$$

Саме цю нерівність звичайно називають *нерівністю Чебишева.*
7) Для будь-якого $\varepsilon > 0$ є справедливою оцінка

$$P\{|\xi| > \varepsilon\} \le \frac{M|\xi|}{\varepsilon}\,. \tag{5.26}$$

$\triangleleft$ Використаємо (5.20) та (5.24). Тоді $p = P\{|\xi| \le \varepsilon\}\,, \quad q = 1 - p = P\{|\xi| > \varepsilon\}$. Оскільки $|\eta| \le \xi$, то $M|\xi| \ge M|\eta| = 0 \cdot p + \varepsilon \cdot q = \varepsilon P\{|\xi| > \varepsilon\}$, що доводить (5.26). $\triangleright$
8) Дисперсія постійної величини дорівнює нулеві:

$$Dc = 0\,. \tag{5.27}$$

9) Обернене твердження: Якщо $D\xi = 0$, то з ймовірністю 1 величина ξ є сталою: $\xi = M\xi$.

$\triangleleft$ Дійсно, внаслідок (5.25) для будь-якого $\varepsilon > 0$ маємо $P\{|\xi - M\xi| > \varepsilon\} = 0$. Тоді

$$\begin{aligned} P\{|\xi - M\xi| > 0\} = & P\{|\xi - M\xi| > 1\} + P\left\{\tfrac{1}{2} < |\xi - M\xi| \le 1\right\} + \\ & + \ldots + P\left\{\frac{1}{2^k} < |\xi - M\xi| \le \frac{1}{2^{k-1}}\right\} + \ldots = 0\,. \end{aligned}$$

У тому числі $\lim\limits_{k\to\infty} P\left\{\dfrac{1}{2^k} < |\xi - M\xi|\right\} = 0$. Тоді

$$P\{\xi = M\xi\} = 1\,.$$

$\triangleright$

10) Якщо $\eta = c\xi$, то

$$D\eta = c^2 D\xi\,. \tag{5.28}$$

11) Дисперсія суми попарно незалежних доданків дорівнює сумі дисперсій

$$D\left(\sum_{i=1}^{n}\xi_i\right)=\sum_{i=1}^{n}D\xi_i\,. \tag{5.29}$$

$\triangleleft$

$$\begin{aligned}
D\left(\sum_{i=1}^{n}\xi_1\right) &= M\left(\sum_{i=1}^{n}\xi_i - M\sum_{i=1}^{n}\xi_i\right)^2 = M\left(\sum_{i=1}^{n}(\xi_i - M\xi_i)\right)^2 = \\
&= M\sum_{i,k=1}^{n}(\xi_i-M\xi_i)(\xi_k-M\xi_k)=\sum_{i,k=1}^{n}M\left[(\xi_i-M\xi_i)(\xi_k-M\xi_k)\right]= \\
&= \sum_{i=1}^{n}M(\xi_i-M\xi_i)^2+\sum_{i\neq k}M\left[(\xi_i-M\xi_i)(\xi_k-M\xi_k)\right]=\textstyle\sum_{i=1}^{n}D\xi_i\,.
\end{aligned}$$

$\triangleright$

Приклад 36. *Нехай випадкова величина ξ розподілена за біноміальним законом. Знайти $M\xi$ та $D\xi$.*

$\triangleleft$ Нехай ξ_k дорівнює кількості успіхів в k-му випробуванні в серії із n незалежних випробувань Бернуллі. Нехай ймовірність успіху при кожному випробуванні дорівнює p, тобто ξ_k приймає два значення (1 з ймовірностю $P\{\xi_k=1\}=p$ та 0 з ймовірностю $P\{\xi_k=0\}=1-p=q$). Тоді

$$M\xi_k=1\cdot p+0\cdot q=p\,,\; D\xi_k=M\xi^2-(M\xi_k)^2=1\cdot p+0\cdot q-p^2=p(1-p)=pq\,.$$

Кількість успіхів ξ у серії з n випробувань дорівнює сумі $\xi=\xi_1+\ldots+\xi_n$. Тоді, внаслідок незалежності ξ_i та враховуючи (5.16) та (5.29), маємо

$$M\xi=\sum_{i=1}^{n}M\xi_i=np\,,\quad D\xi=\sum_{i=1}^{n}D\xi_i=npq\,. \tag{5.30}$$

$\triangleright$

5.3. Умовне математичне сподівання

Раніше ми ввели умовну функцію розподілу $F(x|B)=P\{\xi<x|B\}$ випадкової величини ξ при умові B, якщо $P(B)>0$.
Означення. Математичне сподівання ξ відносно до цього умовного розподілу називається *умовним математичним сподіванням*:

$$M(\xi|B)=\int_{-\infty}^{\infty}x\,dF_\xi(x|B) \tag{5.31}$$

та

$$M(\xi|B)=\sum_{k=1}^{\infty}x_kp_k(B)\,, \tag{5.32}$$

де $p_k(B) = P\{\xi = x_k|B\}$. Очевидно, що (5.31) можна переписати у вигляді

$$M(\xi|B) = \int\limits_{-\infty}^{\infty} x\, p_\xi(x|B)dx\,, \tag{5.33}$$

де $p_\xi(x|B)$ - умовна густина ймовірності. Зокрема, якщо подія B полягає в тому, що випадкова величина η приймає значення y, тобто $\eta = y$, то

$$M(\xi|y) \equiv M(\xi|\eta = y) = \int\limits_{-\infty}^{\infty} x\, p_\xi(x|y)dx\,. \tag{5.34}$$

Аналогічний вираз має місце для дискретних випадкових величин.

Визначені таким чином умовні математичні сподівання мають властивості 1)-7) звичайних математичних сподівань, але в них є деякі додаткові специфічні властивості, що пов'язані з можливістю застосування до них різних варіантів формули повної ймовірності. Тому отримані нижче формули широко застосовуються в теорії ймовірностей і математичній статистиці.

Якщо є повна група попарно несумісних подій B_k, $k = 1, 2, \cdots, n$ і $F_\xi(x|B_k)$ - відповідні умовні функції розподілу, то можна записати

$$F_\xi(x) = \sum_{k=1}^{n} F_\xi(x|B_k)P(B_k)\,. \tag{5.35}$$

Звідси отримаємо

$$M\xi = \sum_{k=1}^{n} M(\xi|B_k)P(B_k)\,. \tag{5.36}$$

Очевидно, що формулу (5.36) можна розглядати як математичне сподівання від нової випадкової величини (дискретної), що приймає значення $M(\xi|B_k)$ з ймовірністю $P(B_k)$. Тобто (5.36) можна переписати у вигляді

$$M\xi = M\left\{M(\xi|B_k)\right\}\,. \tag{5.37}$$

Якщо подія B_k полягає у тому, що випадкова величина η приймає значення y, то природно записати

$$M\xi = \int\limits_{\infty}^{\infty} M(\xi|y)p_\eta(y)dy\,. \tag{5.38}$$

Аналогічна формула має місце і в дискретному випадку.

Для доведення (5.38) помножимо (5.34) на $P_\eta(y)$ і проінтегруємо від $-\infty$ до ∞

$$\int\limits_{-\infty}^{\infty} dy\, M(\xi|y)p_\eta(y) = \int\limits_{-\infty}^{\infty} dy \int\limits_{-\infty}^{\infty} dx\, x\, p_\xi(x|y)p_\eta(y) =$$

$$= \int\limits_{-\infty}^{\infty} dx \cdot x \cdot \int\limits_{-\infty}^{\infty} dy \cdot p(x,y) = \int\limits_{-\infty}^{\infty} dx \cdot x \cdot p_\xi(x) = M\xi\,.$$

Очевидно, що у цій рівності величину $M(\xi|y)$, що стоїть під знаком інтеграла, можна розглядати як функцію від випадкової величини η і позначити її $M(\xi|\eta)$. Тоді (5.38) еквівалентно запису

$$M\xi = M\left\{M(\xi|\eta)\right\}\,. \tag{5.39}$$

Умовне математичне сподівання $M(\xi|\eta)$, що розглядається як функція η, у статистиці називається *функцією регресії* величини ξ на η. Якщо, наприклад,

$$M(\xi|\eta) = \alpha_1\eta + \alpha_2\,, \tag{5.40}$$

то має місце *лінійна регресія*, а α_1 та α_2 - *коефіцієнти регресії.*

5.4. Моменти векторних випадкових величин

Означення. Математичним сподіванням вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ називається вектор

$$M\vec{\xi} = (M\xi_1, \ldots, M\xi_n)\,. \tag{5.41}$$

Означення. Дисперсією вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ називається вектор

$$D\vec{\xi} = (D\xi_1, \ldots, D\xi_n)\,. \tag{5.42}$$

Для багатовимірних величин також справедлива теорема про математичне сподівання функції від випадкової величини (див. рівн. (5.7)), тобто, якщо $\eta = f(\vec{\xi})$, то

$$M\eta = Mf(\vec{\xi}) = \int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_n)p(x_1, \ldots, x_n)dx_1 \ldots dx_n \tag{5.43}$$

за умови, що (5.43) збігається абсолютно. Зокрема, для скалярного добутку

$$M(\vec{a}\cdot\vec{\xi}) = M\sum_{k=1}^{n} a_k\xi_k = \sum_{k=1}^{n} a_k\cdot M\xi_k = \left(\vec{a}\cdot M\vec{\xi}\right)\,. \tag{5.44}$$

Якщо координати випадкового вектора незалежні у сукупності, то

$$M(\xi_1\cdot\xi_2\cdots\xi_n) = M(\xi_1)\cdot M(\xi_2)\cdots M(\xi_n)\,. \tag{5.45}$$

Означення. Коваріаційною (дисперсійною) матрицею називається матриця з елементами

$$a_{ij} = \operatorname{cov}\xi_i\cdot\xi_j \equiv M\left[(\xi_i - M\xi_i)(\xi_j - M\xi_j)\right] = M\xi_i\xi_j - M\xi_i M\xi_j\,.$$

Тут на головній діагоналі стоять дисперсії $a_{ii} = D\xi_i$. Недіагональний елемент також називається *кореляційним моментом* $a_{ij} = a_{ji}$. Очевидно, що, якщо ξ_i та ξ_j незалежні, то $a_{ij} = \operatorname{cov}\xi_i\xi_j = 0$. Таким чином, умова $a_{ij} \neq 0$ є достатньою ознакою залежності ξ_i та ξ_j.

Обернене твердження є невірним, тобто з рівності нулеві $\operatorname{cov}\xi_i\xi_j$ не випливає незалежність ξ_i та ξ_j.

Для будь-якого c маємо $D(\xi_i + c\xi_j) = D\xi_i + c^2 D\xi_j + 2c \cdot \operatorname{cov}\xi_i\xi_j \geq 0$. Виберемо c у вигляді $c = -\dfrac{\operatorname{cov}\xi_i\xi_j}{D\xi_j}$. Тоді $D\xi_j - \dfrac{(\operatorname{cov}\xi_i\xi_j)^2}{D\xi_j} \geq 0$, тобто

$$|\operatorname{cov}\xi_i\xi_j| \leq \sqrt{D\xi \cdot D\xi_j}\,. \tag{5.46}$$

Властивість (5.29) перетворюється

$$D\left(\sum_{i=1}^{n}\xi_i\right) = \sum_{i=1}^{n} D\xi_i + \sum_{\substack{i,j=1\\ i\neq j}}^{n} \operatorname{cov}\xi_i\xi_j\,. \tag{5.47}$$

Для характеристики чистого (лінійного) зв'язку між ξ_i та ξ_j вводиться т.зв. *коефіцієнт кореляцій*

$$r_{ij} = \frac{\operatorname{cov}\xi_i\xi_j}{\sqrt{D\xi \cdot D\xi_j}}\,. \tag{5.48}$$

Внаслідок (5.46) має місце обмеження $-1 \leq r_{ij} \leq +1$. Використовуючи (5.47), отримаємо ($\sqrt{D\xi_i} = \sigma_i$)

$$D\left(\frac{\xi_i}{\sigma_i} \pm \frac{\xi_j}{\sigma_j}\right) = 1 + 1 \pm 2r_{ij} = 2(1 \pm r_{ij}) \geq 0\,. \tag{5.49}$$

Звідси випливає, що $r_{ij} = \pm 1$ тоді і тільки тоді, коли $D(\eta) = D\left(\dfrac{\xi_i}{\sigma_i} \pm \dfrac{\xi_j}{\sigma_j}\right) = 0$. А це можливо якщо тільки $\eta = \left(\dfrac{\xi_i}{\sigma_i} \pm \dfrac{\xi_j}{\sigma_j}\right)$ - стала величина. Таким чином, якщо $r_{ij} = \pm 1$, то ξ_i та ξ_j зв'язані лінійно

$$\xi_i = \alpha\xi_j + \beta\,.$$

У загальному випадку, якщо $r_{ij} > 0$, то кореляція додатна, тобто ξ_i та ξ_j зростають і спадають одночасно. Якщо $r_{ij} < 0$, то кореляція від'ємна (*антикореляція*). Якщо $r_{ij} = 0$, то ξ_i та ξ_j *некорельовані.*

Приклад 37. *χ_n^2-розподілом (розподілом Пірсона) з n степенями свободи називається розподіл випадкової величини $\chi_n^2 = \xi_1^2 + \ldots + \xi_n^2$, де усі $\xi_i \in N(0,1)$ і незалежні.*

$\triangleleft$ Випадковий вектор $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ має густину

$$p(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right]\,. \tag{5.50}$$

Тоді, відповідно до (4.57) та (4.58) маємо

$$F(x)=P\{\chi_n^2<x\}=\underset{\sum x_i^2<x}{\int\dots\int}\frac{1}{(2\pi)^{n/2}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}x_i^2\right]dx_1\dots dx_n\,.$$

Визначимо сферичні координати

$$\left.\begin{array}{lcl} x_1 & = & \rho\cos\varphi_1 \\ x_2 & = & \rho\sin\varphi_1\cos\varphi_2 \\ \vdots & & \\ x_{n-1} & = & \rho\sin\varphi_1\sin\varphi_2\dots\cos\varphi_{n-1} \\ x_n & = & \rho\sin\varphi_1\sin\varphi_2\dots\sin\varphi_{n-1} \end{array}\right\}$$

$$-\tfrac{\pi}{2}\le\varphi_i\le\tfrac{\pi}{2},\quad i=1,2,\dots,n-2,\quad -\pi\le\varphi_{n-1}\le\pi\,.$$

У нових змінних функція розподілу має вигляд (J - якобіан переходу):

$$\begin{aligned} F(x) &= \frac{1}{(2\pi)^{n/2}}\int\dots\int d\rho d\varphi_1\dots d\varphi_{n-1}J(\rho,\varphi_1,\dots,\varphi_{n-1})\cdot \\ &\cdot\exp\left(-\frac{1}{2}\rho^2\right)=\frac{1}{(2\pi)^{n/2}}\int\limits_0^{\sqrt{x}}d\rho\cdot\rho^{n-1}\cdot\exp\left(-\frac{1}{2}\rho^2\right)\cdot \\ &\cdot\int\limits_{-\pi/2}^{\pi/2}\dots\int\limits_{-\pi/2}^{\pi/2}\int\limits_{-\pi}^{\pi}d\varphi_1\dots d\varphi_{n-1}J^{(1)}(\varphi_1,\dots,\varphi_{n-1})= \\ &=\frac{\omega_{n-1}}{(2\pi)^{n/2}}\int\limits_0^{\sqrt{x}}d\rho\cdot\rho^{n-1}\cdot\exp\left(-\frac{1}{2}\rho^2\right). \end{aligned}$$

Тут $J^{(1)}$ - визначник матриці, яка відрізняється від матриці в J відсутністю множника ρ в $n-1$ стовпчиках; ω_{n-1} - інтеграл від цього визначника $J^{(1)}$ по кутових змінних. Сталу ω_{n-1} легко знайти з умови

$$F(\infty)=\frac{\omega_{n-1}}{(2\pi)^{n/2}}\int\limits_0^{\infty}d\rho\cdot\rho^{n-1}\cdot\exp\left(-\frac{1}{2}\rho^2\right)=\frac{\omega_{n-1}}{(2\pi)^{n/2}}\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}-1}=1\,.$$

Тобто

$$\omega_{n-1}=\frac{2\cdot\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}$$

і

$$F(x)=\frac{1}{2^{n/2-1}\Gamma\left(\frac{n}{2}\right)}\int\limits_0^{\sqrt{x}}d\rho\cdot\rho^{n-1}\cdot\exp\left(-\frac{1}{2}\rho^2\right),\quad x>0\,. \tag{5.51}$$

Оскільки $F(x) = 0$ при $x \le 0$, то

$$p_{\chi_n^2}(x) = F'(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0\,, \\ 0, & x \le 0\,. \end{cases} \tag{5.52}$$

Крім цього

$$M\chi_n^2 = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} \int\limits_0^\infty dx\, x^{\frac{n}{2}}\, e^{-\frac{x}{2}} = \frac{2^{\frac{n}{2}+1}}{2^{n/2}\Gamma(\frac{n}{2})} \Gamma\left(\frac{n}{2}+1\right) = n\,,$$

а також

$$\begin{aligned} D\chi_n^2 &= M\left(\chi_n^2\right)^2 - \left(M\chi_n^2\right)^2 = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} \int\limits_0^\infty dx\, x^{\frac{n}{2}+1}\, e^{-\frac{x}{2}} - n^2 = \\ &= \frac{2^{\frac{n}{2}+2}}{2^{n/2}\Gamma(\frac{n}{2})} \Gamma\left(\tfrac{n}{2}+2\right) - n^2 = n^2 + 2n - n^2 = 2n\,. \end{aligned}$$

$\triangleright$

Приклад 38. *t_n-розподілом (розподілом Ст'юдента) з n степенями свободи називаєтьcя розподіл випадкової величини $t_n = \xi/\eta$, де $\xi \in N(0,1)$, $\eta = \sqrt{\xi_n^2/n}$, ξ_n^2 розподілена за Пірсоном, ξ та η незалежні.*

$\triangleleft$ Відповідно до цього $p(x_1, x_2) = p_\xi(x_1)p_\eta(x_2)$ і використаємо (4.64)

$$p_{t_n}(z) = F'_{t_n}(z) = -\int\limits_{-\infty}^0 dy \cdot y \cdot p_\xi(zy)p_\eta(y) + \int\limits_0^\infty dy \cdot y \cdot p_\xi(zy)p_\eta(y)\,. \tag{5.53}$$

Розглянемо

$$F_\eta(x) = P\left\{\sqrt{\frac{\chi_n^2}{n}} < x\right\} = P\left\{\chi_n^2 < nx^2\right\} = F_{\chi_n^2}(nx^2)\,.$$

Тоді, відповідно до (5.52), маємо

$$\begin{aligned} p_{\chi_n^2}(x) &= F'_\eta(x) = p_{\chi_n^2}(nx^2) 2nx = \\ &= \begin{cases} \dfrac{n^{\frac{n}{2}}}{2^{n/2-1}\Gamma(\frac{n}{2})} x^{n-1} e^{-\frac{nx^2}{2}}, & x > 0\,, \\ 0, & x \le 0\,. \end{cases} \end{aligned} \tag{5.54}$$

Внаслідок (5.54) перший доданок у (5.53) зникає, і знаходимо

$$\begin{aligned} p_{t_n}(z) &= \int\limits_0^\infty dy \cdot y \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2y^2}{2}} \cdot \frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})} y^{n-1} e^{-\frac{ny^2}{2}} = \\ &= \frac{n^{\frac{n}{2}} 2^{\frac{n}{2}-1}}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})\sqrt{2\pi}} \cdot \frac{1}{(z^2+n)^{\frac{n}{2}+\frac{1}{2}}} \int\limits_0^\infty dt \cdot t^{\frac{n-1}{2}} \cdot e^{-t} = \\ &= \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{\pi n}} \left(1 + \frac{z^2}{n}\right)^{-\frac{n+1}{2}}\,. \end{aligned}$$

Математичне сподівання Mt_n $(n \geq 2)$

$$Mt_n = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{\pi n}} \int\limits_{-\infty}^{\infty} x \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} dx = 0\,, \tag{5.55}$$

оскільки підінтегральна функція непарна. Дисперсія

$$Dt_n = Mt_n^2 = M\left(\frac{\xi^2}{\chi^2/n}\right) = M\xi^2 \cdot M\left(\frac{n}{\chi_n^2}\right)\,, \tag{5.56}$$

оскільки ξ^2 та χ_n^2/n незалежні. Позначимо $\eta = n/\chi_n^2$, тоді $(x > 0)$

$$F_\eta(x) = P\left\{\frac{n}{\chi_n^2} < x\right\} = P\left\{\frac{1}{x} < \frac{\chi_n^2}{n}\right\} = 1 - P\left\{\chi_n^2 \leq \frac{n}{x}\right\}\,.$$

Використаємо (5.51) із заміною $x \to \dfrac{n}{x}$

$$p_\eta(x) = \begin{cases} \left(\dfrac{n}{2}\right)^{\frac{n}{2}} \dfrac{1}{\Gamma(\frac{n}{2})} e^{-\frac{n}{2x}} x^{-(\frac{n}{2}+1)}, & x > 0\,, \\ 0, & x \leq 0\,. \end{cases} \tag{5.57}$$

Тоді

$$\begin{aligned} M\left(\tfrac{n}{\chi_n^2}\right) &= M\eta = \left(\frac{n}{2}\right)^{\frac{n}{2}} \frac{1}{\Gamma\left(\frac{n}{2}\right)} \int\limits_0^\infty x^{-\frac{n}{2}} e^{-\frac{n}{2x}} dx = \\ &= \frac{1}{\Gamma\left(\frac{n}{2}\right)} \cdot \frac{n}{2} \cdot \int\limits_0^\infty t^{\frac{n}{2}-1} e^{-t} dt = \frac{n}{n-2}\,. \end{aligned}$$

І оскільки $M\xi^2 = D\xi = 1$, то

$$Dt_n = \frac{n}{n-2}\,. \tag{5.58}$$

Розділ 6

Закони великих чисел

Означення. Послідовність випадкових величин $\{\xi_n\}$ *прямує за ймовірністю* до випадкової величини ξ, якщо для будь-якого $\varepsilon > 0$

$$\lim_{n\to\infty} P\{|\xi_n - \xi| > \varepsilon\} = 0\,, \tag{6.1}$$

і позначається

$$\xi_n \xrightarrow{p} \xi,\; n \to \infty \quad \text{або} \quad p\text{-}\lim_{n\to\infty} \xi_n = \xi\,. \tag{6.2}$$

Теорема 6. *Нехай $p\text{-}\lim\limits_{n\to\infty} \xi_n = \xi$ і $g(x)$ є неперервною функцією, $x \in R_1$, так що $\eta = g(\xi)$ та $\eta_n = g(\xi_n)$ є випадковими величинами ($n = 1, 2, \ldots$). Тоді $p\text{-}\lim\limits_{n\to\infty} \eta_n = \eta$.*

Доведення не приводимо.

Лема 1. *Якщо для послідовності випадкових величин $\{\xi_n\}$, $M\xi_n = 0$, $D\xi_n = 0$ при $n \to \infty$, то $p\text{-}\lim\limits_{n\to\infty} \xi_n = 0$.*

$\lhd$ Внаслідок нерівності Чебишева (5.25) і того, що $M\xi_n = 0$, маємо для довільного $\varepsilon > 0$ співвідношення

$$0 \le P\{|\xi_n| > \varepsilon\} \le \frac{D\xi_n}{\varepsilon^2} \longrightarrow 0,\; n \to \infty\,, \tag{6.3}$$

тобто $p\text{-}\lim\limits_{n\to\infty} \xi_n = 0$. $\rhd$

Теорема 7. (Чебишева - Закон великих чисел) *Нехай $\xi_1, \xi_2, \ldots$ - послідовність випадкових попарно незалежних величин, дисперсія яких обмежена у сукупності: $D\xi_n \le c$. Тоді послідовність випадкових величин $\eta_n = \dfrac{1}{n}\sum\limits_{i=1}^{n}(\xi_i - M\xi_i)$ прямує за ймовірністю до нуля при $n \to \infty$:*

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_i \xi_i - \frac{1}{n}\sum_i M\xi_i\right| > \varepsilon\right\} = 0\,, \quad \varepsilon > 0\,. \tag{6.4}$$

$\triangleleft$ Маємо $M\eta_n = \frac{1}{n}\sum\limits_{i=1}^{n}(M\xi_i - M\xi_i) = 0$, і внаслідок попарної незалежності

$$D\eta_n = \frac{1}{n^2}\sum_{i=1}^{n} D\xi_i \le \frac{cn}{n^2} \stackrel{n\to\infty}{\longrightarrow} 0\,.$$

Тоді, виходячи з наведеної вище леми, випливає твердження теореми. $\triangleright$
Рівняння (6.4) можна переписати інакше

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_i \xi_i - \frac{1}{n}\sum_i M\xi_i\right| \le \varepsilon\right\} = 1\,, \quad \varepsilon > 0 \quad \text{- довільне}\,.$$

Наслідок 7.1. *Якщо* $M\xi_1 = M\xi_2 = \ldots = \mu$, *то послідовність* $\eta_n = \dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ *при* $n \to \infty$ *прямує за ймовірністю до математичного сподівання* μ*:*

$$\eta_n = \frac{1}{n}\sum_{i=1}^{n}\xi_i \stackrel{p}{\longrightarrow} \mu, \quad n \to \infty\,. \tag{6.5}$$

Цей наслідок теореми Чебишева є обґрунтуванням правила середнього арифметичного. Якщо при вимірах відсутня систематична похибка (тобто усі $M\xi_i = \mu$), то згідно з законом великих чисел при достатньо великих n із ймовірністю близької до одиниці буде отриманий результат $\widehat{\mu}$, що як завгодно мало відрізняється від істинного значення μ.

Теорема 8. (Бернуллі) *Нехай* η_n *- кількість успіхів у серії із* n *випробувань Бернуллі і* p *- ймовірність успіху при кожному випробуванні. Тоді послідовність частот* $\{\eta_n/n\}$ *при* $n \to \infty$ *прямує за ймовірністю до* p*.*

$\triangleleft$ Нехай ξ_k - кількість успіхів при k-ому випробуванні, $k = 1, 2, \ldots$, тобто $\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n$, $M\xi_k = p$, $D\xi_k = pq$. Тоді відповідно до теореми Чебишева для будь-якого $\varepsilon > 0$ має місце рівність

$$\lim_{n\to\infty} P\left\{\left|\frac{\eta}{n} - p\right| > \varepsilon\right\} = \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_k \xi_k - \frac{1}{n}\sum_k M\xi_k\right| > \varepsilon\right\} = 0\,.$$

$\triangleright$

Означення. Послідовність випадкових величин $\{\xi_n\}$ прямує до випадкової величини ξ з *ймовірністю 1* (або *майже напевно*), якщо

$$P\left\{\omega \in \Omega : \lim_{n\to\infty}\xi_n(\omega) = \xi(\omega)\right\} = 1\,. \tag{6.6}$$

Ця збіжність позначається $\xi_n \to \xi$ м.н.

Лема 2. (Бореля-Кантеллі) *Якщо для послідовності подій* $\{A_n\}$, $n{=}1, 2\ldots$, *виконується умова*

$$\sum_{n=1}^{\infty} P\{A_n\} < \infty\,, \tag{6.7}$$

то з ймовірністю 1 *відбудеться лише скінченна кількість цих подій.*

$\triangleleft$ Нехай подія B_n полягає в тому, що відбувається хоча б одна з подій A_k із $k \geq n$, тобто $B_n = \bigcup_{k=n}^{\infty} A_k$. Очевидно, що $B_1 \supset B_2 \supset \ldots$ Нехай подія B означає, що відбувається нескінчена кількість подій із A_n, $n = 1, 2, \ldots$ Подія B відбувається тоді і тільки тоді, коли відбуваються усі B_n, тобто $B = \bigcap_{n=1}^{\infty} B_n$. Тоді внаслідок неперервності ймовірності

$$P(B) = \lim_{n\to\infty} P\left\{B_n\right\} . \tag{6.8}$$

Оскільки $B_n = \bigcup_{k=n}^{\infty} A_k$, то

$$P(B_n) \leq \sum_{k=n}^{\infty} P\left(A_k\right) . \tag{6.9}$$

Внаслідок збіжності ряду (6.7) його залишок $\sum\limits_{k=n}^{\infty} P\left(A_k\right) \to 0$ при $n \to \infty$, і внаслідок (6.9) $P(B_n) \to 0$ при $n \to \infty$. Тоді із (6.8) випливає $P(B) = 0$ і протилежна подія $\overline{B}$, що складається з того, що відбувається тільки скінченна кількість подій A_n, має ймовірність $P(\overline{B}) = 1$. $\triangleright$

Теорема 9. (Посилений закон великих чисел.) *Нехай $\xi_1, \xi_2, \ldots$ - послідовність попарно незалежних випадкових величин, для яких $M\xi_i = \mu$, $D\xi_i = \sigma^2$. Тоді при $n \to \infty$ має місце*

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \longrightarrow \mu \tag{6.10}$$

з ймовірністю 1.

$\triangleleft$ Введемо нову випадкову величину $\xi_i' = \xi_i - \mu$, тобто можна вважати $\mu = 0$. Нехай $\eta_k = \sum\limits_{i=1}^{k} \xi_i$. Нам треба показати, що при $n \to \infty$ величина $\dfrac{1}{n}\eta_n \to 0$ м.н. Для кожного натурального n візьмемо натуральне m таким чином, щоб

$$m^2 \leq n \leq (m+1)^2 . \tag{6.11}$$

Оскільки $M\eta_k = 0$, то нерівність Чебишева (5.25) дає оцінку

$$P\left\{\left|\frac{\eta_{m^2}}{m^2}\right| > \varepsilon\right\} \leq \frac{D\eta_{m^2}}{\varepsilon^2 m^4} = \frac{\sigma^2}{\varepsilon^2 m^4} . \tag{6.12}$$

Покладемо $\widehat{\eta}_{m^2} = \max\limits_{m^2+1\leq k\leq(m+1)^2} \left|\xi_{m^2} + \ldots + \xi_k\right|$. Застосуємо нерівність Чебишева

$$\begin{aligned} &P\left\{\left|\tfrac{\widehat{\eta}_{m^2}}{m^2}\right| > \varepsilon\right\} \leq \sum_{k=m^2+1}^{(m+1)^2} P\left\{\frac{\left|\xi_{m^2} + \ldots + \xi_k\right|}{m^2} > \varepsilon\right\} \leq \\ &\leq \sum_{k=m^2+1}^{(m+1)^2} \frac{(k-m^2)\sigma^2}{\varepsilon^2 m^4} \leq 2m\frac{(2m+1)\sigma^2}{\varepsilon^2 m^4} \leq \frac{5\sigma^2}{\varepsilon^2 m^4} . \end{aligned} \tag{6.13}$$

Внаслідок (6.12) та (6.13) числові ряди

$$\sum_{m=1}^{\infty} P\left\{\left|\frac{\eta_{m^2}}{m^2}\right| > \varepsilon\right\} \quad \text{та} \quad \sum_{m=1}^{\infty} P\left\{\left|\frac{\widehat{\eta}_{m^2}}{m^2}\right| > \varepsilon\right\}$$

збігаються. Тоді виходячи із леми Бореля-Кантеллі з ймовірністю 1 відбувається тільки скінченна кількість подій $\left\{\left|\frac{\eta_{m^2}}{m^2}\right| > \varepsilon\right\}$ та $\left\{\left|\frac{\widehat{\eta}_{m^2}}{m^2}\right| > \varepsilon\right\}$, тобто з ймовірністю 1

$$\frac{\eta_{m^2}}{m^2} \to 0 \quad \text{та} \quad \frac{\widehat{\eta}_{m^2}}{m^2} \to 0 \quad \text{при} \quad m \to \infty\,. \tag{6.14}$$

Оскільки для довільного n із співвідношення (6.11) має місце оцінка

$$\left|\frac{\eta_n}{n}\right| \le \left|\frac{\eta_{m^2}}{m^2}\right| + \left|\frac{\widehat{\eta}_{m^2}}{m^2}\right|\,,$$

то з (6.14) випливає, що $\frac{\eta_n}{n} \to 0$ при $n \to \infty$ з ймовірністю 1. $\triangleright$

Теорема 10. (Бореля) *Нехай η_n - кількість успіхів у серії із n незалежних випробувань Бернуллі, p - ймовірність успіху при кожному випробуванні. Тоді послідовність частот $\left\{\frac{\eta_n}{n}\right\}$ при $n \to \infty$ з ймовірністю 1 прямує до ймовірності p.*

$\triangleleft$ Нехай ξ_i - кількість успіхів в i-ому випробуванні ($\xi_i = 0, 1$; $m\xi_i = p$, $D\xi_i = pq$). Введемо величину $\eta_n = \sum_i \xi_i$ і використаємо попередню теорему. $\triangleright$

Теорема 11. (Хінчина) *Нехай однаково розподілені випадкові величини $\xi_1, \xi_2, \ldots$ попарно незалежні і мають скінченні математичні сподівання $M\xi_i = \mu$. Тоді при $n \to \infty$ послідовність $\eta_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k$ збігається за ймовірністю до μ: $p\text{-}\lim_{n\to\infty} \eta_n = \mu$.*

$\triangleleft$ Введемо послідовність нових випадкових величин $\overline{\xi}_k$ за формулою

$$\overline{\xi}_k = \begin{cases} \xi_k, & \text{якщо} \quad |\xi_k| \le y\,, \\ 0, & \text{якщо} \quad |\xi_k| > y\,, \end{cases}$$

де y - довільно. Розглянемо випадкову величину $\overline{\eta}_n = \frac{1}{n}\sum_{k=1}^{n} \overline{\xi}_k$. Тоді для довільного $\varepsilon > 0$ маємо ($\overline{\mu} = M\overline{\xi}_k$)

$$P\left\{|\eta_n - \overline{\mu}| > \varepsilon\right\} \le P\left\{|\overline{\eta}_n - \overline{\mu}| > \varepsilon\right\} + P\left\{\eta_n \ne \overline{\eta}_n\right\}\,, \tag{6.15}$$

оскільки, якщо відбувається подія в лівій частині, то відбувається хоча б одна подія в правій частині. Далі

$$\overline{\mu} = M\overline{\xi}_k = \int_{-y}^{y} x\,dF(x)\,.$$

Тому

$$|\overline{\mu} - \mu| < \varepsilon \,, \tag{6.16}$$

якщо $y \geq y_0 = y_0(\varepsilon)$. Тоді внаслідок (6.15) та (6.16)

$$P\left\{|\eta_n - \overline{\mu}| > 2\varepsilon\right\} \leq P\left\{|\overline{\eta}_n - \overline{\mu}| > \varepsilon\right\} + P\left\{\eta_n \neq \overline{\eta}_n\right\} \,.$$

Відповідно до (5.25) маємо

$$P\left\{|\overline{\eta}_n - \overline{\mu}| > \varepsilon\right\} \leq \frac{D\overline{\eta}_n}{\varepsilon^2} \leq \frac{1}{\varepsilon^2} M\overline{\eta}_n^2 \leq \frac{1}{\varepsilon^2 n^2} \sum_{k=1}^{n} M\overline{\xi}_k^2 =$$
$$= \frac{n}{\varepsilon^2 n^2} \int\limits_{-y}^{y} x^2 dF(x) \leq \frac{y}{\varepsilon^2 n} \int\limits_{-y}^{y} |x| dF(x) \leq \frac{y}{\varepsilon^2 n} \int\limits_{-\infty}^{\infty} |x| dF(x) = \frac{Ay}{\varepsilon^2 n} \,,$$

де $A = \int_{-\infty}^{\infty} |x| dF(x)$. Подія $\{\eta_n \neq \overline{\eta}_n\}$ полягає в тому, що хоча б для однієї випадкової величини ξ_k виконано $|\xi_k| > \varepsilon$, $k = 1, \ldots, n$. Тому для будь-якого $\varepsilon_1 > 0$

$$P\left\{\eta_n \neq \overline{\eta}_n\right\} \leq \sum_{k=1}^{n} P\left\{|\xi_k| > y\right\} = n \int\limits_{|x|>y} dF(x) \leq n \frac{1}{y} \int\limits_{|x|>y} |x| dF(x) \leq \frac{n}{y} \varepsilon_2$$

за умови, що $y \geq y_1 = y_1(\varepsilon_1)$. Тоді при $y \geq \overline{y} = \max\{y_0, y_1\}$ є справедливою оцінка

$$P\left\{|\eta_n - \mu| > 2\varepsilon\right\} \leq \frac{Ay}{\varepsilon^2 n} + \frac{n}{y} \varepsilon_1 \,.$$

Для будь-якого $\delta > 0$ знайдуться $y = \dfrac{\varepsilon^2}{A} \delta n$ та $\varepsilon_1 = \dfrac{\delta^2}{\varepsilon^2 A}$. І тоді

$$P\left\{|\eta_n - \mu| > 2\varepsilon\right\} \leq 2\delta \,.$$

$\triangleright$

Розділ 7

Центральні граничні теореми

Закони великих чисел містять у собі висновки про збіжність послідовностей випадкових величин $\{\xi_n,\ n = 1, 2, \ldots\}$ до деякої випадкової (або невипадкової) величини ξ. Ці твердження не дають нам ніякої інформації про те, як апроксимувати розподіл випадкових величин ξ_n при великих n. Відповідь на це питання дають центральні граничні теореми, де вводиться нове поняття - збіжність за розподілом.

7.1. Характеристична функція

Означення. Характеристичною функцією (х.ф.) випадкової величини ξ називається функція $f_\xi(t)$ дійсної змінної $t \in R_1$

$$f_\xi(t) = Me^{it\xi} = \int\limits_{-\infty}^{\infty} e^{itx} dF_\xi(x)\,. \tag{7.1}$$

У загальному випадку, коли $\chi = \xi + i\eta$ - комплексна випадкова величина, то за означенням

$$M\chi = M\xi + iM\eta\,. \tag{7.2}$$

У виразі (7.1) інтеграл необхідно розуміти або як ряд, що абсолютно збігається

$$f_\xi(t) = \sum_{k=1}^{\infty} e^{itx} p_k\,, \tag{7.3}$$

або як інтеграл, що абсолютно збігається

$$f_\xi(t) = \int\limits_{-\infty}^{\infty} e^{itx} p(x) dx\,. \tag{7.4}$$

Х.ф. існує для довільної випадкової величини, оскільки внаслідок рівності $|e^{itx}| = 1$ ряд (7.3) та інтеграл (7.4) збігаються абсолютно.

Очевидно, що $f_\xi(0) = 1$ та $|f_\xi(t)| \le 1$, $t \in R_1$.

Теорема 12. *Нехай $\xi_1, \ldots, \xi_n$ - незалежні у сукупності випадкові величини. Тоді*

$$f_{\xi_1+\ldots+\xi_n}(t) = f_{\xi_1}(t) \cdots f_{\xi_n}(t)\,. \tag{7.5}$$

$\lhd$

$$\begin{aligned} f_{\xi_1+\ldots+\xi_n}(t) &= Me^{it(\xi_1+\ldots+\xi_n)} = M\left(e^{it\xi_1}\cdots e^{it\xi_n}\right) = \\ &= Me^{it\xi_1}\cdots Me^{it\xi_n} = f_{\xi_1}(t)\cdots f_{\xi_n}(t)\,. \end{aligned}$$

$\rhd$

Теорема 13. *Якщо σ та μ - деякі сталі, то*

$$f_{\sigma\xi+\mu}(t) = e^{it\mu} f_\xi(\sigma t)\,. \tag{7.6}$$

$\lhd$

$$f_{\sigma\xi+\mu}(t) = Me^{it(\sigma\xi+\mu)} = e^{it\mu}Me^{it\sigma\xi} = e^{it\mu}f_\xi(\sigma t)\,. \tag{7.7}$$

$\rhd$

Теорема 14. *Якщо існує момент $M\xi^k$, то k-а похідна $f_\xi^{(k)}(t)$ від х.ф. $f_\xi(t)$ існує, рівномірно неперервна на R_1 та*

$$f_\xi^{(k)}(0) = i^k M\xi^k\,. \tag{7.8}$$

$\lhd$ За означенням

$$f_\xi^{(k)}(0) = \int\limits_{-\infty}^{\infty} e^{itx}(ix)^k dF_\xi(x)\Big|_{t=0} = i^k M\xi^k\,.$$

Рівномірну неперервність доводити не будемо. $\rhd$

Теорема 15. *Якщо х.ф. $f_\xi(t)$ абсолютно інтегрована на R_1, то випадкова величина ξ неперервна, а її густина $p(x)$ дорівнює*

$$p(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itx} f_\xi(t)dt\,, \tag{7.9}$$

і є рівномірно неперервною на R_1.

$\lhd$ Нехай маємо два довільні числа $x < y$.Розглянемо інтеграл

$$j(x,y) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\frac{e^{-itx}-e^{-ity}}{it}f_\xi(t)dt = \frac{1}{2\pi}\lim_{A\to\infty}\int\limits_{-A}^{A}\frac{e^{-itx}-e^{-ity}}{it}f_\xi(t)dt\,. \tag{7.10}$$

Оскільки

$$\left|e^{-itx}-e^{-ity}\right| = \left|\int\limits_{ty}^{tx} e^{-i\alpha}d\alpha\right| \le |t(x-y)|\,, \tag{7.11}$$

і $f_\xi(t)$ є абсолютно інтегрованою, то (7.10) збігається абсолютно. Із (7.11) також випливає, що

$$j(x, x+h) \to 0 \quad \text{при} \quad h \to 0\,. \tag{7.12}$$

Підставимо (7.1) в (7.10) і змінимо порядок інтегрування

$$\begin{aligned} j(x,y) &= \frac{1}{2\pi}\lim_{A\to\infty}\int\limits_{-\infty}^{\infty}\left[\int\limits_{-A}^{A}\frac{e^{it(z-x)}-e^{it(z-y)}}{it}dt\right]dF_\xi(z) = \\ &=\frac{1}{\pi}\int\limits_{-\infty}^{\infty}\lim_{A\to\infty}dF_\xi(z)\int\limits_{0}^{A}\left[\frac{e^{it(z-x)}-e^{-it(z-x)}}{2i}-\frac{e^{it(z-y)}-e^{-it(z-y)}}{2i}\right]\frac{dt}{t} = \\ &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty}dF_\xi(z)\left[\int\limits_{0}^{\infty}\frac{\sin(z-x)t}{t}dt-\int\limits_{0}^{\infty}\frac{\sin(z-y)t}{t}dt\right]. \end{aligned} \tag{7.13}$$

Використаємо формулу $\int\limits_0^\infty \frac{\sin\alpha t}{t}dt = \frac{\pi}{2}\operatorname{sign}\alpha$. Тоді вираз у квадратних дужках в останньому виразі рівності (7.13) дорівнює нулеві, якщо $z < x$ або $z > y$, та дорівнює π, якщо $x < z < y$. Тоді

$$j(x,y) = \int\limits_x^y dF_\xi(z)\,. \tag{7.14}$$

Із (7.12) випливає, що $F_\xi(x)$ - неперервна функція по x, а із (7.14) випливає, що

$$j(x,y) = F_\xi(y) - F_\xi(x)\,. \tag{7.15}$$

Звідси можна знайти

$$\begin{aligned} &\frac{F_\xi(x+h)-F_\xi(x-h)}{2h} = \frac{j(x-h,x+h)}{2h} = \\ &= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\frac{e^{-it(x-h)}-e^{-it(x+h)}}{2ith}f_\xi(t)dt = \\ &=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\frac{e^{ith}-e^{-ith}}{2ith}e^{-itx}f_\xi(t)dt=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\frac{\sin th}{th}e^{-itx}f_\xi(t)dt \overset{h\to 0}{\Longrightarrow} \\ &\Longrightarrow \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}e^{-itx}f_\xi(t)dt = F'_\xi(x)\,. \end{aligned}$$

Тобто

$$p(x) = F'_\xi(t) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}e^{-itx}f_\xi(t)dt\,. \tag{7.16}$$

Рівномірну збіжність доводити не будемо. $\triangleright$
Формули (7.10), (7.15) та (7.16) називаються *формулами обернення*, вони дозволяють знаходити густини і функції розподілу за відомими х.ф.

Приклад 39. *Х.ф. нормального розподілу $N(0,1)$ має вигляд*

$$f_\xi(t) = \int\limits_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = e^{-\dfrac{t^2}{2}} . \tag{7.17}$$

Якщо $\xi \in N(0,1)$, то $\eta = (\sigma\xi + \mu) \in N(\mu, \sigma^2)$ і відповідно до (7.7) маємо

$$f_\eta(t) = e^{it\mu - \dfrac{\sigma^2 t^2}{2}} . \tag{7.18}$$

Приклад 40. *Х.ф. розподілу Пуассона має вигляд*

$$f_\xi(t) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \left(\lambda e^{it}\right)^k \frac{1}{k!} = e^{\lambda e^{it} - \lambda} . \tag{7.19}$$

Приклад 41. *Х.ф. біноміального розподілу має вигляд*

$$f_\xi(t) = \sum_{k=0}^{n} e^{itk} C_n^k p^k q^{n-k} = \sum_{k=0}^{n} C_n^k \left(pe^{it}\right)^k q^{n-k} = \left(pe^{it} + q\right)^n . \tag{7.20}$$

Означення. Характеристичною функцією $f_{\vec{\xi}}(t_1, \ldots, t_n)$, $(t_1, \ldots, t_n) \in R_n$ n-вимірної випадкової величини $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ називається функція

$$f_{\vec{\xi}}(t_1, ., t_n) = Me^{it_1\xi_1 + . + it_n\xi_n} = \int\limits_{-\infty}^{\infty} . \int \exp\left[i \sum_{k=1}^{n} t_k x_k \right] dF_{\vec{\xi}}(x_1, \ldots, x_n)) .$$

Звідси випливають основні властивості х.ф.

$$1) \qquad f_{\vec{\xi}}(\vec{0}) = f_{\vec{\xi}}(0, \ldots, 0) = 1, \ \left| f_{\vec{\xi}}(\vec{t}) \right| = \left| f_{\vec{\xi}}(t_1, \ldots, t_n) \right| \le 1 .$$

2) Якщо $\xi_1, \ldots, \xi_n$незалежні, то

$$f_{\vec{\xi}}(\vec{t}) = Me^{i(\vec{\xi}, \vec{t})} = f_{\xi_1}(t_1) \cdots f_{\xi_n}(t_n) . \tag{7.21}$$

3) Х.ф. випадкової величини $\vec{\eta} = (\sigma_1\xi_1 + \mu_1, \ldots, \sigma_n\xi_n + \mu_n)$ дорівнює

$$f_{\vec{\eta}}(t_1, \ldots, t_n) = \exp\left[i \sum_k \mu_k t_k \right] f_{\vec{\xi}}(\sigma_1 t_1, \ldots, \sigma_n t_n) .$$

4) Х.ф. суми координат випадкової величини ξ дорівнює

$$f_{\xi_1+\ldots+\xi_n}(t) = Me^{it(\xi_1+\ldots+\xi_n)} = f_{\vec{\xi}}(t,\ldots,t)\,. \tag{7.22}$$

Із властивостей характеристичної функції (формула (7.8)) випливає, що якщо існують усі моменти випадкової величини ξ, то відповідна х.ф. $f_\xi(t)$ буде нескінченно диференційованою. Такі функції називаються *фінітними*, і ми будемо позначати $f_\xi(t) \in C_0^\infty(R_1)$. Для фінітних функцій є справедливими формули прямого і оберненого перетворення Фур'є: якщо $\varphi(x) \in C_0^\infty(R_1)$, то

$$\varphi(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{itx}\tilde{\varphi}(t)dt \quad \text{та} \quad \tilde{\varphi}(t) = \int\limits_{-\infty}^{\infty} e^{-itx}\varphi(x)dx\,.$$

Причому перетворення Фур'є $\tilde{\varphi}(t)$ спадає на нескінченності швидше будь-якої степені і тому є абсолютно інтегрованим на R_1.

Лема 3. *Нехай $\varphi(x)$ - довільна фінітна функція, $\varphi(x) \in C_0^\infty(R_1)$, ξ - випадкова величина, $f_\xi(t)$ - її х.ф. Тоді*

$$M\varphi(\xi) = \int\limits_{-\infty}^{\infty} f_\xi(t)\tilde{\varphi}(t)dt \tag{7.23}$$

$$\triangleleft \quad M\varphi(\xi) = \int\limits_{-\infty}^{\infty}\varphi(x)dF_\xi(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} dF_\xi(x)\left[\int\limits_{-\infty}^{\infty} e^{itx}\tilde{\varphi}(t)dt\right] =$$
$$= \quad \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} dt\tilde{\varphi}(t)\left[\int\limits_{-\infty}^{\infty} dF_\xi(x)e^{itx}\right] = \int\limits_{-\infty}^{\infty} f_\xi(t)\tilde{\varphi}(t)dt\,.$$

Перестановка інтегралів можлива внаслідок абсолютної збіжності подвійного інтеграла. $\triangleright$

Лема 4. *Нехай послідовність х.ф. $\{f_{\xi_n}(t)\}$ випадкових величин $\{\xi_n\}$ збігається при $n\to\infty$ до х.ф. випадкової величини ξ рівномірно по t у кожному обмеженому інтервалі $|t| \le T$. Тоді для будь-якої фінітної функції $\varphi(x) \in C_0^\infty(R_1)$ має місце рівність*

$$M\varphi(\xi_n) \longrightarrow M\varphi(\xi) \quad \textit{при} \quad n\to\infty\,.$$

Доведення не приводимо.

Теорема 16. (Про неперервність характеристичної функції). *Нехай виповнені умови попередньої леми. Нехай функція розподілу $F_\xi(x)$ випадкової величини ξ неперервна. Тоді при $n\to\infty$*

$$\lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)\,, \tag{7.24}$$

причому збіжність рівномірна по $x \in R_1$.

Доведення не приводимо. Є справедливим і обернене твердження.

Теорема 17. *Якщо в кожній точці неперервності $F_\xi(x)$ виконується умова (7.24), то*

$$\lim_{n\to\infty} f_{\xi_n}(t) = f_\xi(t)$$

рівномірно по t.

Приклад 42. *Нехай* $\eta=\xi_1+\xi_2$, *де* $\xi_1 \in N(\mu_1, \sigma_1^2)$, $\xi_2 \in N(\mu_2, \sigma_2^2)$ *і незалежні. Знайти розподіл* η.

$\lhd$ Відповідно до (7.18), (7.21) і (7.22) маємо

$$f_{\xi_1+\xi_2}(t)=f_\eta(t)=f_{\xi_1}(t)f_{\xi_2}(t)=\exp\left[it(\mu_1+\mu_2)-\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}\right].$$

Звідки випливає $\eta \in N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$, тобто

$$p_\eta(x)=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}\exp\left[-\frac{(x-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right]. \rhd$$

Приклад 43. *Нехай* $\eta=\xi_1+\xi_2$, *де* ξ_1 *та* ξ_2 *розподілені за Пуассоном із параметрами* λ_1 *та* λ_2, *відповідно, і незалежні. Знайти розподіл* η.

$\lhd$ Відповідно до (7.19), (7.21) і (7.22) маємо

$$f_{\xi_1+\xi_2}(t)=f_\eta(t)=f_{\xi_1}(t)f_{\xi_2}(t)=\exp\left[\lambda_1\left(e^{it}-1\right)\right]\exp\left[\lambda_2\left(e^{it}-1\right)\right]=$$
$$=\exp\left[(\lambda_1+\lambda_2)\left(e^{it}-1\right)\right].$$

Звідки випливає η розподілена за Пуассоном з параметром $\lambda=\lambda_1+\lambda_2$, тобто

$$p_\eta(k)=\frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}.$$

$\rhd$

Приклад 44. *Знайти густину ймовірностей для* χ_n^2*-розподілу:* $\chi_n^2=\xi_1^2+\ldots+\xi_n^2$, $\xi_i \in N(0,1)$.

$\lhd$ Спочатку отримаємо явний вигляд для густини розподілу одного доданку ξ_i^2:

$$F_{\xi^2}(z)=P\{\xi^2<z\}=\int\limits_{x^2<z}p_\xi(x)dx=\int\limits_{-\sqrt{z}}^{\sqrt{z}}p_\xi(x)dx\,,\ z>0\,,$$

$$p_{\xi^2}(z)=F'_{\xi^2}(z)=\frac{1}{2\sqrt{z}}\left[p_\xi(\sqrt{z})+p_\xi(-\sqrt{z})\right]=\frac{1}{\sqrt{2\pi}\sqrt{z}}\exp\left(-\frac{z}{2}\right).$$

Тоді відповідно до (7.4) та (7.21) маємо

$$f_{\chi_n^2}(t)=f_{\xi_1^2}(t)\cdots f_{\xi_1^2}(t)=\left[\int\limits_0^\infty e^{itx}\frac{1}{\sqrt{2\pi}\sqrt{x}}\exp\left(-\frac{x}{2}\right)dx\right]^n=$$
$$=\left[\frac{1}{\sqrt{2\pi}}\frac{\sqrt{2\pi}}{\sqrt{1-2it}}\right]^n=(1-2it)^{-\dfrac{n}{2}},$$

і внаслідок (7.16)

$$p_{\chi_n^2}(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}e^{-itx}(1-2it)^{-n/2}dt=\frac{1}{2^{\frac{n}{2}}\Gamma(\dfrac{n}{2})}x^{\dfrac{n}{2}-1}\exp\left[-\frac{x}{2}\right],$$

що співпадає з отриманим раніше результатом (5.52). $\rhd$

7.2. Центральні граничні теореми

Означення. Послідовність випадкових величин $\{\xi_k\}$, $k = 1, 2, \ldots$ прямує до випадкової величини ξ_0 *за розподілом* або *слабо збігається*, якщо

$$\lim_{k\to\infty} F_k(x) = F_0(x) \tag{7.25}$$

у кожній точці неперервності $F_0(x)$, де $F_k(x)$ - функція розподілу ξ_k.

Теорема 18. (Центральна гранична теорема для однаково розподілених незалежних випадкових величин). *Нехай $\{\xi_k\}$ - послідовність незалежних однаково розподілених випадкових величин із $M\xi_k = \mu$, $D\xi_k = \sigma^2$, $\sigma > 0$, $k = 1, 2, \ldots$. Тоді послідовність випадкових величин*

$$\eta_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (\xi_k - \mu) \tag{7.26}$$

збігається за розподілом до $N(0, 1)$, тобто

$$\lim_{n\to\infty} P\{\eta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{z^2}{2}\right) dz = \Phi_0(x)\,, \tag{7.27}$$

де $\Phi_0(x)$ є функцією розподілу нормально розподіленої величини $N(0, 1)$. При цьому прямування до границі в (7.27) є рівномірним по $x \in R_1$.

$\triangleleft$ Введемо позначення

$$\chi_k = \xi_k - \mu\,,\ M\chi_k = 0\,,\ \eta_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} \chi_k\,.$$

Внаслідок (7.6) та однакового розподілу для усіх χ_k маємо

$$f_{\eta_n}(t) = \left[\varphi_{\chi_k}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n = \left[\varphi_{\chi_k}(0) + \varphi'_{\chi_k}(0)\frac{t}{\sigma\sqrt{n}} + \varphi''_{\chi_k}(0)\frac{t^2}{2\sigma^2 n} + o\left(\frac{t^2}{2\sigma^2 n}\right)\right]^n,$$

де залишковий член має форму Пеано. Тут $\varphi_{\chi_k}(t)$ є характеристичною функцією випадкової величини χ_k. Врахуємо формулу (7.8) і те, що $\varphi_{\chi_k}(0) = 1$, $M\chi_k = 0$, $M\chi_k^2 = \sigma^2$. Тоді

$$f_{\eta_n}(t) = \left[1 - \frac{t^2}{2\sigma^2 n} + o\left(\frac{t^2}{2\sigma^2 n}\right)\right]^n \longrightarrow_{n\to\infty} \exp\left(-\frac{t^2}{2}\right)$$

рівномірно по t. $\triangleright$

Наслідок 18.2. (Інтегральна теорема Муавра-Лапласа) *Нехай ξ - кількість успіхів у серії із n незалежних іспитів, p - ймовірність успіху при кожному іспиті. Тоді при $n \to \infty$*

$$\lim_{n\to\infty} P\left\{\frac{\xi - np}{\sqrt{npq}} < x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{z^2}{2}\right] dz\,,$$

причому прямування до границі є рівномірним по $x \in R_1$.

$\lhd$ Маємо $\xi = \xi_1 + \ldots + \xi_n$, де ξ_k - кількість успіхів при k-ому іспиті, ξ_k незалежні і однаково розподілені, $P\{\xi_k = 1\} = p$, $P\{\xi_k = 0\} = q$, $M\xi_k = p$, $D\xi_k = pq$. Щоб порівняти із (7.26) необхідно покласти

$$\xi_k \Longleftrightarrow \xi = \xi_1 + \ldots + \xi_n,\ M\xi_k = \mu \Longleftrightarrow M\xi = np\,.$$

Теорему доведено. $\rhd$

Теорема 19. *Якщо поряд з умовами попередньої теореми х. ф. випадкової величини ξ_k є абсолютно інтегрованою на R_1, то густина $p_n(x)$ випадкової величини (7.26) при $n \to \infty$ збігається до густини $p_0(x) = \dfrac{1}{\sqrt{2\pi}} \exp\left[-\dfrac{x^2}{2}\right]$ рівномірно по x.*

Доведення не приводимо.

Зараз ми позбавимося від небажаного у багатьох випадках припущення про однаковий розподіл випадкових величин ξ_k.

Теорема 20. (Центральна гранична теорема Ляпунова) *Нехай $\{\xi_k\}$ – послідовність незалежних випадкових величин, що мають $M\xi_k = \mu_k$, $D\xi_k = \sigma_k^2$ та $M|\xi_k - \mu_k|^3 < \infty$, $k = 1, 2, \ldots$ Тоді, якщо виконується умова Ляпунова*

$$\frac{1}{B_n^3} \sum_{k=1}^{n} M|\xi_k - \mu_k|^3 \longrightarrow 0 \qquad \textit{при} \qquad n \to \infty\,,$$

де $B_n = \sqrt{\sum\limits_{k=1}^{n} D\xi_k} = \sqrt{\sum\limits_{k=1}^{n} \sigma_k^2}$, то послідовність випадкових величин

$$\eta_n = \frac{1}{B_n} \sum_{k=1}^{n} (\xi_k - \mu_k)$$

збігається за розподілом до $N(0,1)$ рівномірно на R_1.

$\lhd$ Покладемо $\chi_{nk} = \dfrac{\xi_k - \mu_k}{B_n}$. Тоді

$$\eta_n = \sum_{k=1}^{n} \chi_{nk}\,, \qquad M\chi_{nk} = \frac{1}{B_n} M(\xi_k - \chi_k) = 0\,,$$
$$D\chi_{nk} = \frac{1}{B_n^2} D\xi_k = \frac{\sigma_k^2}{B_n^2}\,, \quad M|\chi_{nk}|^3 = \frac{1}{B_n^3} M|\xi_k - \mu_k|^3\,.$$

Внаслідок незалежності для х.ф. маємо співвідношення

$$f_{\eta_n}(t) = \prod_{k=1}^{n} \varphi_{\chi_{nk}}(t)\,, \tag{7.28}$$

де $\varphi_{\chi_{nk}}(t)$ - х.ф. випадкової величини χ_{nk}.

Розкладемо $\varphi_{\chi_{nk}}(t) \equiv \varphi(t)$ в ряд Тейлора із залишковим членом у формі Лагранжа

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \varphi''(0)\frac{t^2}{2} + R_{nk}(t) = 1 - \frac{\sigma_k^2}{B_n^2}\frac{t^2}{2} + \varphi'''_{\chi_{nk}}(\theta)\frac{t^3}{6}\,.$$

При $n \to \infty$ оцінка цього залишкового члена проводиться за допомогою співвідношення (7.8), тобто

$$\left|\varphi'''_{\chi_{nk}}(\theta)\right| \leq \frac{1}{B_n^3}M|\xi_k - \mu_k|^3 \longrightarrow 0\,.$$

Тоді

$$\ln \varphi_{\chi_{nk}}(t) = -\frac{\sigma_k^2}{B_n^2} \cdot \frac{t^2}{2} + O(t^3)\,,$$

і внаслідок (7.28) маємо

$$\ln f_{\eta_n}(t) = -\frac{\sum\limits_k \sigma_k^2}{B_n^2} \cdot \frac{t^2}{2} + O(t^3) \longrightarrow_{n\to\infty} -\frac{t^2}{2}$$

та

$$f_{\eta_n}(t) = \exp\left[-\frac{t^2}{2}\right].$$

Рівномірну збіжність доводити не будемо. ▷

7.3. Застосування центральних граничних теорем

Спочатку зазначимо випадки, коли не можна застосовувати ЦГТ. Нехай ми хочемо оцінити ймовірність $P\{\eta_n < x\}$, причому мова йде про ті значення x, при яких ця ймовірність близька до нуля або одиниці. Якщо ми замінимо $P\{\eta_n < x\} \to \Phi_0(x)$, то відносна похибка може бути дуже великою, оскільки, хоча різниця $P\{\eta_n < x\} - \Phi_0(x)$ рівномірна мала по x, але невірно твердження, що відношення $P\{\eta_n < x\}/\Phi_0(x) \to 1$ рівномірно по x, тобто "хвости"розподілів потребують дуже обережної оцінки.

Оцінимо, наскільки сильно може відрізнятися частота від ймовірності у серії з n випробувань Бернуллі. Оцінка ґрунтується на співвідношенні

$$P\left\{\left|\frac{\eta_n}{n} - p\right| > \varepsilon\right\} = P\left\{\left|\frac{\eta_n - np}{\sqrt{npq}}\right| > \varepsilon\sqrt{\frac{n}{pq}}\right\} \approx$$

$$\approx \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{-\varepsilon\sqrt{\frac{n}{pq}}} e^{-\frac{x^2}{2}}dx + \frac{1}{\sqrt{2\pi}}\int\limits_{\varepsilon\sqrt{\frac{n}{pq}}}^{\infty} e^{-\frac{x^2}{2}}dx = 2\Phi_0\left(-\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Оскільки $p + q = 1$, то $pq = p(1-p) = \frac{1}{4} - \frac{1}{4} + p - p^2 = \frac{1}{4} - (\frac{1}{2} - p)^2$, тобто $pq + (\frac{1}{2} - p)^2 = \frac{1}{4}$, звідки $pq \leq \frac{1}{4}$. Тоді

$$2\Phi_0\left(-\varepsilon\sqrt{\frac{n}{pq}}\right) \leq 2\Phi_0\left(-2\varepsilon\sqrt{n}\right)\,.$$

Звідси має місце оцінка

$$P\left\{\frac{\eta_n}{n}-\varepsilon<p<\frac{\eta_n}{n}+\varepsilon\right\}\approx 1-2\Phi_0\left(-\varepsilon\sqrt{\frac{n}{pq}}\right)\geq 1-2\Phi_0\left(-2\varepsilon\sqrt{n}\right).$$

Таким чином, якщо відома кількість успіхів в серії з n випробувань Бернуллі, можна побудувати інтервал $(\eta_n/n-\varepsilon, \eta_n/n+\varepsilon)$, який буде містити істинне (невідоме) значення ймовірності p з довільною заданою ймовірністю $1-\alpha$. Для цього просто необхідно за таблицями обрати значення ε таким, щоб $2\Phi_0(-2\varepsilon\sqrt{n})=\alpha$. Тоді

$$P\left\{\frac{\eta_n}{n}-\varepsilon(\alpha)<p<\frac{\eta_n}{n}+\varepsilon(\alpha)\right\}=1-\alpha\,. \tag{7.29}$$

Інтервал $\left(\frac{\eta_n}{n}-\varepsilon(\alpha), \frac{\eta_n}{n}+\varepsilon(\alpha)\right)$ називається *довірчим інтервалом з рівнем довіри* $1-\alpha$.

Нехай $\xi_1, \xi_2, \ldots$ - незалежні випадкові величини, $M\xi_k=\mu$, $D\xi_k=\sigma^2$ для усіх k. Закон великих чисел (6.5) стверджує, що при $n\to\infty$ є справедливою рівність

$$P\left\{\left|\frac{\xi_1+\ldots+\xi_n}{n}-\mu\right|>\varepsilon\right\}=0\,.$$

Якщо $\xi_1, \xi_2, \ldots$ не тільки попарно незалежні (що є достатнім для застосування закону великих чисел), але й незалежні у сукупності, то можна застосувати теорему (7.27):

$$P\left\{\left|\frac{\xi_1+\ldots+\xi_n}{n}-\mu\right|>\varepsilon\right\}=P\left\{\left|\frac{\xi_1+\ldots+\xi_n-n\mu}{\sigma\sqrt{n}}\right|>\frac{\varepsilon\sqrt{n}}{\sigma}\right\}=$$
$$=P\left\{\eta_n<-\frac{\varepsilon\sqrt{n}}{\sigma}\right\}+P\left\{\eta_n>\frac{\varepsilon\sqrt{n}}{\sigma}\right\}=2\Phi_0\left(-\frac{\varepsilon\sqrt{n}}{\sigma}\right).$$

Із таблиць випливає, що при $\frac{\varepsilon\sqrt{n}}{\sigma}=3$ (тобто $\varepsilon=\frac{3\sigma}{\sqrt{n}}$) ймовірність

$$P\left\{|\eta_n|<\frac{\varepsilon\sqrt{n}}{\sigma}\right\}=1-2\Phi_0\left(-\frac{\varepsilon\sqrt{n}}{\sigma}\right) \tag{7.30}$$

дорівнює 0.997. Це є *правило* 3 σ.

Розділ 8

Математична статистика

Теорія ймовірностей вивчає математичні моделі експериментів із випадковим результатом. Математична статистика вивчає питання, що виникають при порівнянні цих моделей із реальністю.

8.1. Розподіл ортогональних проекцій

Нехай $\vec{\xi}$ - випадковий вектор в n-вимірному евклідовому просторі R_n з базисом $\{\vec{e_i}\}$: $\vec{\xi} = \sum\limits_{i=1}^{n} \xi_i \vec{e_i}$. Розглянемо *скалярний добуток* у цьому просторі $\left(\vec{\xi}, \vec{\eta}\right) = \sum\limits_{i=1}^{n} \xi_i \eta_i$ Величини $\xi_i = \left(\vec{e_i}, \vec{\xi}\right)$ (проекції вектора) є випадковими величинами.

Нехай A - лінійний оператор в R_n. Тоді $\vec{\eta} = A\vec{\xi}$ є випадковим вектором, причому

$$\eta_i = (\vec{e_i}, \vec{\eta}) = \left(\vec{e_i}, A\vec{\xi}\right) = \sum_{j=1}^{n} a_{ij} \xi_j \,, \tag{8.1}$$

де величини $a_{ij} = (\vec{e_i}, A\vec{e_j})$ є матричними елементами оператора A в ортономованому базисі $\{\vec{e_i}\}$. Відповідну матрицю будемо позначати $\breve{A}$. Перейдемо до нового ортонормованого базису $\{\vec{e'_i}\}$ за допомогою *ортогонального перетворення* U

$$\vec{e'_i} = U\vec{e_i}, \quad i = 1, \ldots, n \,,$$

яке має такі властивості ($\breve{U}$ - матриця перетворення U)

$$\breve{U}^T = \breve{U}^{-1}, \quad \det \breve{U} = \pm 1 \,. \tag{8.2}$$

У новому базисі матричні елементи a'_{ij} оператора A обчислюються за формулами

$$a'_{ij} = \left(\vec{e'_i}, A\vec{e'_j}\right) = (U\vec{e_i}, AU\vec{e_j}) = \left(\vec{e_i}, U^+ AU\vec{e_j}\right) = \sum_{k,l=1}^{n} u^T_{ik} a_{kl} u^T_{lj} \,.$$

Нагадаємо, що внаслідок (8.2) $(\vec{x}, \vec{y}) = (U\vec{x}, U\vec{y})$ і норма вектора зберігається $\|\vec{x}\| = \|U\vec{x}\|$, де $\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})}$.

Нехай L - лінійний підпростір R_n і $L^\perp$ - ортогональне доповнення L в R_n, тобто

$$L^\perp = \{\vec{x} \in R_n,\, (\vec{x},\vec{y}) = 0,\, \vec{y} \in L\}\ . \tag{8.3}$$

Покажемо, що розклад

$$\vec{x} = \vec{x_1} + \vec{x_2},\, \vec{x_1} \in L,\, \vec{x_2} \in L^\perp \tag{8.4}$$

є единим. Дійсно, із рівності $\vec{x} = \vec{x'_1} + \vec{x'_2}$, $\vec{x'_1} \in L$, $\vec{x'_2} \in L^\perp$ і лінійності підпросторів L та $L^\perp$ випливає $(\vec{x_1} - \vec{x'_1}) + (\vec{x_2} - \vec{x'_2}) = 0$. Оскільки $\Big((\vec{x_1} - \vec{x'_1}), (\vec{x_2} - \vec{x'_2})\Big)$, то $\|\vec{x_1} - \vec{x'_1}\|^2 + \|\vec{x_2} - \vec{x'_2}\|^2 = 0 \implies \vec{x_1} = \vec{x'_1}$, $\vec{x_2} = \vec{x'_2}$. Таким чином, кожному вектору $\vec{x} \in R_n$ розклад (8.4) ставить у відповідність вектор $\vec{x_1} \in L$, або

$$\vec{x_1} = \Pi\vec{x}\,, \tag{8.5}$$

де Π - *оператор проектування на L, (ортогональний проектор).*

Властивості оператора проектування Π:

1) Π - лінійний оператор. Дійсно, нехай

$$\vec{x} = \vec{x_1} + \vec{x_2},\, \vec{y} = \vec{y_1} + \vec{y_2},\, \vec{x_1}, \vec{y_1} \in L,\, \vec{x_2}, \vec{y_2}, \in L^\perp\,.$$

Тоді

$$\alpha\vec{x} + \beta\vec{y} = (\alpha\vec{x_1} + \beta\vec{y_1}) + (\alpha\vec{x_2} + \beta\vec{y_2}),\, \alpha\vec{x_1}, \beta\vec{y_1} \in L,\, \alpha\vec{x_2}, \beta\vec{y_2}, \in L^\perp\,.$$

Як наслідок

$$\Pi(\alpha\vec{x} + \beta\vec{y}) = \alpha\vec{x_1} + \beta\vec{y_1} = \alpha\Pi\vec{x} + \beta\Pi\vec{y}. \tag{8.6}$$

2) Π - самоспряжений оператор, тобто для будь-яких $\vec{x}, \vec{y} \in R_n$

$$(\Pi\vec{x}, \vec{y}) = (\vec{x}, \Pi\vec{y})\ . \tag{8.7}$$

Дійсно

$$(\Pi\vec{x}, \vec{y}) = (\vec{x_1}, \vec{y}) = (\vec{x_1}, \vec{y_1}) = (\vec{x}, \vec{y_1}) = (\vec{x}, \Pi\vec{y})\ .$$

3) $$\Pi^2 = \Pi\,. \tag{8.8}$$

Дійсно, для довільного $\vec{x} \in R_n$ має місце

$$\Pi\vec{x} = \vec{x_1} = \Pi\vec{x_1} = \Pi\,(\Pi\vec{x})\ .$$

Властивості 1)-3) є необхідними і достатніми для того, щоб оператор Π був ортогональним проектором.

Визначимо *відстань* від вектора $\vec{x}$ до L

$$\rho(\vec{x}, L) = \inf\{\|\vec{x} - \vec{y}\|, \vec{y} \in L\}\ . \tag{8.9}$$

Тоді

$$\rho(\vec{x}, L) = \|\vec{x} - \Pi\vec{x}\|\,.$$

Дійсно, для довільного $\vec{y} \in L$ маємо

$$\Pi\vec{x} - \vec{y} \in L \quad \text{та} \quad \vec{x} - \Pi\vec{x} = (I - \Pi)\vec{x} \in L^{\perp}\,.$$

Як наслідок

$$\|\vec{x} - \vec{y}\|^2 = \|\vec{x} - \Pi\vec{x} + \Pi\vec{x} - \vec{y}\|^2 = \|\vec{x} - \Pi\vec{x}\|^2 + \|\Pi\vec{x} - \vec{y}\|^2 \geq \|\vec{x} - \Pi\vec{x}\|^2,$$

причому рівність виконується тільки у випадку $\Pi\vec{x} = \vec{y}$.

Теорема 21. *Нехай $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ випадковий вектор, координати якого незалежні у сукупності і нормальні $N(0, \sigma^2)$. Якщо U - оператор ортогонального перетворення в R_n, то розподіл вектора $\vec{\eta} = U\vec{\xi}$ співпадає з розподілом вектора $\vec{\xi}$.*

$\lhd$ Маємо

$$p_{\vec{\xi}}(\vec{x}) = p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{x_i^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\dfrac{(\vec{x}, \vec{x})}{2\sigma^2}}\,. \tag{8.10}$$

Далі (якщо $\eta = A\xi$)

$$\begin{aligned} \int\limits_D p_{\vec{\eta}}(\vec{y})d\vec{y} &= P\left\{\vec{\eta} \in D\right\} = P\left\{A\vec{\xi} \in D\right\} = P\left\{\vec{\xi} \in A^{-1}D\right\} = \\ &= \int\limits_{A^{-1}D} p_{\vec{\xi}}(\vec{x})d\vec{x} = \int\limits_D p_{\vec{\xi}}(A^{-1}\vec{y})\left|\det A^{-1}\right| d\vec{y}\,. \end{aligned}$$

Звідси

$$p_{\vec{\eta}}(\vec{x}) = p_{\vec{\xi}}(A^{-1}\vec{x})\left|\det A^{-1}\right|\,, \tag{8.11}$$

і якщо $A = U$, то відповідно до (8.10) та (8.11) маємо

$$p_{\vec{\eta}}(\vec{x}) = p_{\vec{\xi}}(U^{-1}\vec{x}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{(U^{-1}\vec{x}, U^{-1}\vec{x})}{2\sigma^2}\right] = p_{\vec{\xi}}(\vec{x})\,. \tag{8.12}$$

Звідси випливає, що координати вектора η незалежні у сукупності та нормальні $N(0, \sigma^2)$. $\rhd$

Очевидно, що твердження теореми можна сформулювати інакше: якщо координати $\xi_1, \ldots, \xi_n$ випадкового вектора $\vec{\xi}$ незалежні у сукупності і нормальні $N(0, \sigma^2)$ у деякому ортонормованому базисі R_n, то вони незалежні у сукупності і нормальні $N(0, \sigma^2)$ і в будь-якому іншому ортонормованому базисі R_n.

Наслідок 21.3. *Нехай R^k - k-вимірний лінійний підпростір R_n, і Π_k - оператор проектування на R^k. Якщо виконані умови теореми (8.12), то випадкова величина*

$$\|\Pi_k\vec{\xi}\|^2 \equiv \left(\Pi_k\vec{\xi}, \vec{\xi}\right) = \left(\Pi_k\vec{\xi}, \Pi_k\vec{\xi}\right) \tag{8.13}$$

розподілена як випадкова величина $\sigma^2\chi_k^2$.

Доведення не приводимо.

Наслідок 21.4. *Нехай* Π_1 *- оператор ортогонального проектування на одновимірний підпростір* R_1*, що натягнутий на одиничний вектор* $\vec{e}$*, тобто* $\Pi_1\vec{\xi} = (\vec{e},\vec{\xi})\vec{e}$*. Якщо випадковий вектор* $\vec{\xi}$ *задовольняє умовам теореми, то випадкова величина*

$$\tau_{n-1} = \frac{(\vec{e},\vec{\xi})}{\left[\|(I-\Pi_1)\vec{\xi}\|^2/(n-1)\right]^{1/2}} \tag{8.14}$$

має розподіл Ст'юдента з $n-1$ *степенями свободи.*

$\lhd$ Нагадаємо, що розподіл Ст'юдента з k степенями свободи має випадкова величина $\tau_k = \dfrac{\eta}{\left[\frac{\chi_k^2}{k}\right]^{1/2}}$, якщо $\eta \in N(0,1)$, $\chi_k^2 \in \chi^2_{\text{розподіл}}$, або якщо $\eta \in N(0,\sigma^2)$ та $\tau_k = \dfrac{\eta}{\left[\frac{\sigma^2\chi_k^2}{k}\right]^{1/2}}$.
Згідно з теоремою (8.12) величина $(\vec{e},\vec{\xi}) \in N(0,\sigma^2)$ і не залежить від $\|(I-\Pi_1)\vec{\xi}\|^2$. Одначасно $\|(I-\Pi_1)\vec{\xi}\|^2 = \sigma^2\chi^2_{n-1}$. $\rhd$.

Розглянемо окремий випадок, коли $\vec{e} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \vec{e}_i$. Тоді координати вектора $\Pi_1\vec{\xi}$ дорівнюють

$$\left(\Pi_1\vec{\xi}\right)_j = \left[(\vec{e},\vec{\xi})\vec{e}\right]_j = \frac{1}{n}\sum_{i=1}^n \xi_i, \quad j=1,\ldots,n\,,$$

тобто всі координати вектора $\Pi_1\vec{\xi}$ однакові. Далі, вектор $(I-\Pi_1)\vec{\xi}$ має координати

$$\left((I-\Pi_1)\vec{\xi}\right)_j = \left(\vec{\xi}-\Pi_1\vec{\xi}\right)_j = \xi_j - \frac{1}{n}\sum_{i=1}^n \xi_i, \quad j=1,\ldots,n\,.$$

Тоді у цьому випадку величина (8.14) набуває вигляду

$$\tau_{n-1} = \frac{(\vec{e},\vec{\xi})}{\left[\|(I-\Pi_1)\vec{\xi}\|^2/(n-1)\right]^{1/2}} = \frac{\frac{1}{\sqrt{n}}\sum\limits_{i=1}^n \xi_i}{\left[\frac{1}{n-1}\sum\limits_{j=1}^n\left(\xi_j - \frac{1}{n}\sum\limits_{i=1}^n \xi_i\right)^2\right]^{1/2}}\,. \tag{8.15}$$

У випадку, коли ξ_i, $i=1,\ldots,n$ незалежні у сукупності і нормальні $N(\mu,\sigma^2)$, то у формулі (8.15) замість $\vec{\xi}$ треба підставити вектор $\vec{\xi}-\vec{m}$, де $\vec{m} = \mu\sum_i \vec{e}_i$. Тоді випадкова величина

$$\tau_{n-1} = \frac{\left(\vec{e},(\vec{\xi}-\vec{m})\right)}{\left[\|(I-\Pi_1)\vec{\xi}\|^2/(n-1)\right]^{1/2}} = \frac{\frac{1}{\sqrt{n}}\sum\limits_{i=1}^n (\xi_i-\mu)}{\left[\frac{1}{n-1}\sum\limits_{j=1}^n\left(\xi_j - \frac{1}{n}\sum\limits_{i=1}^n \xi_i\right)^2\right]^{1/2}} \tag{8.16}$$

має розподіл Ст'юдента з $n-1$ степенями свободи. Тут ми використали той факт, що $(I-\Pi_1)\vec{m} = 0$.

8.2. Оцінка математичного сподівання при відомій дисперсії

Нехай ξ_i, $i=\overline{1,n}$, нормальні $N(\mu,\sigma^2)$, причому дисперсія σ^2 є відомою, а математичне сподівання μ – невідомо. Нехай $\{\xi_i\}$ незалежні у сукупності. Послідовність $\xi_1,\ldots,\xi_n$ називається *випадковою виборкою* із нормального розподілу $N(\mu,\sigma^2)$ об'ємом n.

Оскільки випадкова величина $\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ нормальна $N(\mu,\frac{\sigma^2}{n})$, то випадкова величина

$$\eta=\left(\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\mu)\right)\frac{\sqrt{n}}{\sigma} \tag{8.17}$$

є нормальною $N(0,1)$, і, як наслідок, можна підрахувати ймовірність

$$P\left\{|\eta|<\varepsilon\right\}=\frac{1}{\sqrt{2\pi}}\int\limits_{-\varepsilon}^{\varepsilon}e^{-\frac{x^2}{2}}dx=\Phi_0(\varepsilon)-\Phi_0(-\varepsilon)\,.$$

Задамо значення α, $0<\alpha<1$, і за таблицями нормального розподілу визначимо таке $\varepsilon>0$, для якого

$$P\left\{-\varepsilon<\left(\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\mu)\right)\frac{\sqrt{n}}{\sigma}<\varepsilon\right\}=\Phi_0(\varepsilon)-\Phi_0(-\varepsilon)=1-\alpha\,. \tag{8.18}$$

При цьому випадкова величина η (8.17) може опинитися зовні інтервала $(-\varepsilon,\varepsilon)$ лише із ймовірностю α. Вираз (8.18) можна переписати інакше

$$P\left\{\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{\varepsilon\sigma}{\sqrt{n}}<\mu<\frac{1}{n}\sum_{i=1}^{n}\xi_i+\frac{\varepsilon\sigma}{\sqrt{n}}\right\}=1-\alpha\,. \tag{8.19}$$

Відповідно до (8.19) істинне значення математичного сподівання μ лежить у *випадковому* інтервалі

$$\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{\varepsilon\sigma}{\sqrt{n}}\;,\quad\frac{1}{n}\sum_{i=1}^{n}\xi_i+\frac{\varepsilon\sigma}{\sqrt{n}}\right) \tag{8.20}$$

з ймовірностю $1-\alpha$. Інтервал (8.20) називається *інтервальною оцінкою* параметра μ, або *довірчим інтервалом* для параметра μ (з рівнем довіри $(1-\alpha)\cdot 100\%$).

8.3. Оцінка дисперсії при відомому математичному сподіванні

Нехай маємо ту ж саму послідовність ξ_i із $N(\mu,\sigma^2)$, але μ - відомо, а σ^2 - невідомо. Тоді випадкові величини $(\xi_i-\mu)$, $i=\overline{1,n}$, $\in N(0,\sigma^2)$ і незалежні у сукупності. Як наслідок величина $\sum\limits_{i=1}^{n}(\xi_i-\mu)^2$ має розподіл $\sigma^2\chi_n^2$ з n степенями свободи.

Якщо задані величини $\varepsilon_1 > 0$ та $\varepsilon_2 > 0$, то за таблицями χ^2-розподілу можна підрахувати ймовірність

$$P\left\{\varepsilon_1 < \chi_n^2 < \varepsilon_2\right\} = 1 - \alpha\,. \tag{8.21}$$

Але при заданому α, $0 < \alpha < 1$ величини ε_1 та ε_2 обчислюються неоднозначно. Тому звичайно припускають

$$P\left\{\chi_n^2 \le \varepsilon_1\right\} = P\left\{\chi_n^2 \ge \varepsilon_2\right\} = \frac{\alpha}{2}\,.$$

Тоді виконується умова (8.21) і в результаті отримаємо

$$P\left\{\varepsilon_1 < \frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i - \mu)^2 < \varepsilon_2\right\} = 1 - \alpha\,,$$

або

$$P\left\{\frac{1}{\varepsilon_2}\sum_{i=1}^{n}(\xi_i - \mu)^2 < \sigma^2 < \frac{1}{\varepsilon_1}\sum_{i=1}^{n}(\xi_i - \mu)^2\right\} = 1 - \alpha\,,$$

або

$$P\left\{\frac{1}{\varepsilon_2}\|\vec{\xi} - \vec{m}\|^2 < \sigma^2 < \frac{1}{\varepsilon_1}\|\vec{\xi} - \vec{m}\|^2\right\} = 1 - \alpha\,. \tag{8.22}$$

Відповідно до (8.22) істинне значення дисперсії σ^2 лежить у *випадковому* інтервалі

$$\left(\frac{1}{\varepsilon_2}\|\vec{\xi} - \vec{m}\|^2\,, \quad \frac{1}{\varepsilon_1}\|\vec{\xi} - \vec{m}\|^2\right) \tag{8.23}$$

з ймовірностю $1 - \alpha$. Інтервал (8.23) називається *інтервальною оцінкою* параметра σ^2, або *довірчим інтервалом* для параметра σ^2 (з рівнем довіри $(1 - \alpha) \cdot 100\%$).

8.4. Оцінка математичного сподівання при невідомій дисперсії

Відповідно до (8.16) величина τ_{n-1} має розподіл Ст'юдента з $n - 1$ степенями свободи. За таблицями значень розподілу Ст'юдента для заданого $\alpha \in (0, 1)$ можна визначити ε таке, щоб $P\left\{|\tau_{n-1}| < \varepsilon\right\} = 1 - \alpha$. Або

$$P\left\{\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{\varepsilon\|(I-\Pi_1)\vec{\xi}\|}{\sqrt{n(n-1)}} < \mu < \frac{1}{n}\sum_{i=1}^{n}\xi_i + \frac{\varepsilon\|(I-\Pi_1)\vec{\xi}\|}{\sqrt{n(n-1)}}\right\} = 1 - \alpha\,. \tag{8.24}$$

Зазначимо, що тут

$$\|(I - \Pi_1)\vec{\xi}\| = \sqrt{\sum_{i=1}^{n}\left(\xi_i - \frac{1}{n}\sum_{k=1}^{n}\xi_k\right)^2}\,.$$

Відповідно до (8.24) істинне значення математичного сподівання μ лежить у *випадковому* інтервалі

$$\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{\varepsilon\|(I-\Pi_1)\vec{\xi}\|}{\sqrt{n(n-1)}}\ ,\quad \frac{1}{n}\sum_{i=1}^{n}\xi_i + \frac{\varepsilon\|(I-\Pi_1)\vec{\xi}\|}{\sqrt{n(n-1)}}\right)$$

з ймовірністю $1-\alpha$.

8.5. Оцінка дисперсії при невідомому математичному сподіванні

Відповідно до (8.13) маємо

$$\left\|(I-\Pi_1)(\vec{\xi}-\vec{m})\right\|^2 = \left\|(I-\Pi_1)\vec{\xi}\right\|^2 = \sigma^2\chi^2_{n-1}\,.$$

За заданим $\alpha \in (0,1)1$ визначимо ε_1 та ε_1 за умовами

$$P\left\{\chi^2_{n-1} \le \varepsilon_1\right\} = P\left\{\chi^2_{n-1} \ge \varepsilon_2\right\} = \frac{\alpha}{2}\,.$$

Тоді

$$P\left\{\varepsilon_1 < \frac{1}{\sigma^2}\left\|(I-\Pi_1)\vec{\xi}\right\|^2 < \varepsilon_2\right\} = 1-\alpha\,,$$

або

$$P\left\{\frac{1}{\varepsilon_2}\left\|(I-\Pi_1)\vec{\xi}\right\|^2 < \sigma^2 < \frac{1}{\varepsilon_1}\left\|(I-\Pi_1)\vec{\xi}\right\|^2\right\} = 1-\alpha\,.$$

Випадковий інтервал

$$\left(\frac{1}{\varepsilon_2}\left\|(I-\Pi_1)\vec{\xi}\right\|^2\ ,\quad \frac{1}{\varepsilon_1}\left\|(I-\Pi_1)\vec{\xi}\right\|^2\right) = 1-\alpha \tag{8.25}$$

є *інтервальною оцінкою дисперсії* σ^2 при невідомому математичному сподіванні μ (з рівнем довіри $(1-\alpha)\cdot 100\%$).

8.6. Точкові оцінки

Нехай $\vec{\xi} = (\xi_1,\ldots,\xi_n)$ випадкова вибірка із розподілу $F(x,\vec{\theta})$, що залежить від параметра $\vec{\theta} = (\theta_1,\ldots,\theta_k) \in \Theta$, і нехай $\tau(\vec{\theta})$ - відома функція, що визначена на Θ. Необхідно обчислити $\tau(\vec{\theta})$, але аргумент невідомий. У цьому випадку єдина доступна інформація про $\tau(\vec{\theta})$ міститься у виборці $\vec{\xi}$, і найкраще, що ми можемо зробити, це побудувати оцінку $\tau(\vec{\theta})$, що базується на виборці $\vec{\xi}$. Тобто, необхідно побудувати функцію $t(\cdot)$, що визначена на виборочному просторі R_n, з метою використати статистику $t(\vec{\xi})$ замість $\tau(\vec{\theta})$. Статистика $t(\vec{\xi})$ називається *точковою оцінкою* $\tau(\vec{\theta})$, і ми припускаємо, що розподіл $t(\vec{\xi})$ концентрується біля значення $\tau(\vec{\theta})$, де $\vec{\theta}$ - параметр розподілу вибірки $\vec{\xi}$.

Приклад 45. *Нехай* $\vec{\xi} = (\xi_1,\ldots,\xi_n)$ *є вибіркою із нормального розподілу* $N(\mu,\sigma^2)$, $\vec{\theta} = (\mu,\sigma^2)$, $\Theta = \{-\infty < \mu < \infty, 0 \le \sigma^2\}$. *Розглянемо статистику*

$$\widehat{\mu}_n = t(\vec{\xi}) = \frac{1}{n}\sum_{j=1}^{n}\xi_j\,. \tag{8.26}$$

$\lhd$ Визначимо математичне сподівання $M_{\vec{\theta}}$ і дисперсію $D_{\vec{\theta}}$ в розподілі $F(\cdot, \vec{\theta})$. Маємо

$$M_{\vec{\theta}}\widehat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} M_{\vec{\theta}}\xi_j = \mu\,. \tag{8.27}$$

Тоді, якщо як оцінку математичного сподівання $\mu = \tau(\vec{\theta})$ ми будемо використовувати $\widehat{\mu}_n$, ми не будемо робити систематичної похибки в тому розумінні, що

$$M_{\vec{\theta}}(\widehat{\mu}_n - \mu) = 0\,.$$

Оцінки, що мають властивість (8.27), називаються *незміщеними*.

Далі маємо

$$M_{\vec{\theta}}(\widehat{\mu}_n - \mu)^2 = D_{\vec{\theta}}\widehat{\mu}_n = \frac{1}{n^2}\sum_{j=1}^{n} D_{\vec{\theta}}\xi_j = \frac{\sigma^2}{n}\,.$$

Тобто $D_{\vec{\theta}}\widehat{\mu}_n \to 0$, $n \to \infty$, і відповідно до (6.3) $\widehat{\mu}_n$ за ймовірністю прямує до μ при $n \to \infty$. Це означає, що розподіл $\widehat{\mu}_n$ концентрується навколо μ, оскільки

$$P\{|\widehat{\mu}_n - \mu| > \varepsilon|\vec{\theta}\} \le \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0,\ n \to \infty \tag{8.28}$$

для довільного $\varepsilon > 0$. $\rhd$

Оцінки, що мають властивість (8.28), називаються слушними. Це бажана властивість оцінок, але вона не визначає якості оцінок при фіксованому n. Ми можемо *якість оцінки* визначити як відхилення $\widehat{\mu}_n$ від μ та задати її числом

$$M_{\vec{\theta}}(\widehat{\mu}_n - \mu)^2\,.$$

Природно, чим менше ця величина, тим краще оцінка. Видно, що для незміщеної оцінки ця якість визначається дисперсією. Тому найкращу оцінку μ можна визначити як оцінку із мінімальною дисперсією серед усіх незміщених оцінок.

Дослідимо *роль незміщеності*. Нехай є k незалежних вибірок $\vec{\xi^1} = (\xi_1^1, \ldots, \xi_{n_1}^1)$, …, $\vec{\xi^k} = (\xi_1^k, \ldots, \xi_{n_k}^k)$ із розподілу $F(x, \vec{\theta})$. На основі цих вибірок побудуємо оцінки $t_1(\vec{\xi^1})$, …, $t_k(\vec{\xi^k})$ значення $\tau(\vec{\theta})$, причому

$$M_{\vec{\theta}}t_i(\vec{\xi^i}) = \tau(\vec{\theta}) + \varepsilon_i,\ D_{\vec{\theta}}t_i(\vec{\xi^i}) = \sigma_i^2 \le \sigma^2,\ i = \overline{1,k}\,. \tag{8.29}$$

Побудуємо статистику

$$T_k = \frac{1}{k}\sum_{j=1}^{k} t_j(\vec{\xi^j})\,, \tag{8.30}$$

і знайдемо

$$M_{\vec{\theta}}T_k = \frac{1}{k}\sum_{j=1}^{k} M_{\vec{\theta}}t_j(\vec{\xi^j}) = \tau(\vec{\theta}) + \frac{1}{k}\sum_{j=1}^{k}\varepsilon_i \tag{8.31}$$

та

$$D_{\vec{\theta}}T_k = \frac{1}{k^2}\sum_{j=1}^{k}\sigma_j^2 \le \frac{\sigma^2}{k} \to 0,\ k\to\infty\,. \tag{8.32}$$

Якщо зміщення $\frac{1}{k}\sum\limits_{j=1}^{k}\varepsilon_i$ не прямує до нуля при $k\to\infty$, то T_k при $k\to\infty$ не збігається до справжнього значення $\tau(\vec{\theta})$. Якщо, наприклад, $\varepsilon_j=\varepsilon$, то T_k збігається за ймовірністю до $\tau(\vec{\theta})+\varepsilon$, що є небажаним. З іншого боку, якщо $\varepsilon_j=0$, $j=\overline{1,k}$, то ми отримаємо додавання інформації про $\tau(\vec{\theta})$, що міститься в незалежних незміщених оцінках $t_j(\vec{\xi}^j)$ (тобто зникає інформація про $\tau(\vec{\theta})$).

Разом з тим очевидно, що якщо ми маємо оцінку $\breve{t}$ із мінімальним відхиленням

$$M_{\vec{\theta}}\left(\breve{t}-\tau(\vec{\theta})\right)^2 = \min_t M_{\vec{\theta}}\left(t-\tau(\vec{\theta})\right)^2, \tag{8.33}$$

то за інших рівних умов вимога щодо незміщеності може лише збільшити відхилення (8.33), оскільки умова $M_{\vec{\theta}}t=\tau(\vec{\theta})$ звужує клас оцінок, на яких шукається мінімум в (8.33). Більш того, клас незміщених оцінок взагалі може виявитися порожнім.

Приклад 46. *Нехай ξ - кількість успіхів у послідовності n незалежних іспитів Бернуллі із невідомою ймовірністю $p=\theta$, $0<\theta\le 1$, і нехай необхідно знайти незміщену оцінку $\tau(\theta)=1/\theta$.*

$\lhd$ Для будь-якої оцінки $t(\xi)$ повинна виконуватися рівність

$$M_{\vec{\theta}}t(\xi)=\sum_{i=0}^{n}C_n^i\theta^i(1-\theta)^{n-i}t(i)=\frac{1}{\theta}$$

для всіх $\theta\in(0,1]$. Але це неможливо, оскільки для довільної функції $t(\cdot)$ при $\theta\to 0$ має місце рівність: $M_\theta t(\xi)=C_n^0 t(0)$, а в той же час $\theta^{-1}\to\infty$. $\rhd$

Розглянемо далі статистику

$$\widehat{\sigma}_n^2=\frac{1}{n-1}\sum_{j=0}^{n}(\xi_j-\widehat{\mu}_n)^2=\frac{\|(I-\Pi_1)\vec{\xi}\|^2}{n-1}, \tag{8.34}$$

де $\vec{\xi}=(\xi_1,\dots,\xi_n)$ є вибіркою із $N(\mu,\sigma^2)$. Згідно з (8.13) статистика (8.34) має розподіл $\dfrac{\sigma^2}{n-1}\chi_{n-1}^2$. Тоді

$$\begin{aligned} M_{\vec{\theta}}\widehat{\sigma}_n^2 &= \tfrac{\sigma^2}{n-1}M\chi_{n-1}^2=\sigma^2\,,\\ D_{\vec{\theta}}\widehat{\sigma}_n^2 &= \tfrac{\sigma^4}{(n-1)^2}D\chi_{n-1}^2=\tfrac{2\sigma^4}{n-1}\to 0,\ n\to\infty\,. \end{aligned}$$

Таким чином, $\widehat{\sigma}_n^2$ є незміщеною (слушною) оцінкою $\sigma^2=\tau(\vec{\theta})$, $\vec{\theta}=(\mu,\sigma^2)$. У той же час

$$M_{\vec{\theta}}\left(\frac{1}{n}\|(I-\Pi_1)\vec{\xi}\|^2-\sigma^2\right)^2=\frac{\sigma^4}{n^2}\left[2(n-1)+(n-1)^2\right]-$$
$$-2\frac{\sigma^4}{n}(n-1)+\sigma^4=(\frac{2}{n}-\frac{1}{n^2})\sigma^4<D_{\vec{\theta}}\widehat{\sigma}_n^2,\ n=1,2,\dots$$

Тому для кожного $n = 1, 2, \ldots$ статистика

$$\widetilde{\sigma}_n^2 = \frac{1}{n}\|(I - \Pi_1)\vec{\xi}\|^2$$

менше відхиляється від σ^2, чим $\widehat{\sigma}_n^2$, хоча і має зміщення, оскільки

$$M_{\vec{\theta}}\widetilde{\sigma}_n^2 = \sigma^2 - \frac{\sigma^2}{n}\,.$$

Внаслідок того, що $M_{\vec{\theta}}(\widetilde{\sigma}_n^2 - \sigma^2) = -\dfrac{\sigma^2}{n} \to 0$ при $n \to \infty$, і того, що

$$D_{\vec{\theta}}(\widetilde{\sigma}_n^2 - \sigma^2) = M_{\vec{\theta}}(\widetilde{\sigma}_n^2 - \sigma^2)^2 - \frac{\sigma^4}{n^2} = \frac{2n-2}{n^2} \to 0,\ n \to \infty\,,$$

оцінка $\widetilde{\sigma}_n^2$ збігається за ймовірністю до σ^2, тобто є слушною. В даному випадку зміщеній оцінці дисперсії $\widetilde{\sigma}_n^2$ треба віддати перевагу перед незміщеною $\widehat{\sigma}_n^2$.

Статистики $\widehat{\mu}_n$ (8.26) і $\widehat{\sigma}_n^2$ (8.34) як оцінки для математичного сподівання і дисперсії зберігають ряд властивостей і без припущення про нормальність вибірки. Дійсно, нехай $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ є вибіркою із довільного розподілу, причому

$$M\xi_i = \mu,\ D\xi_i = \sigma^2,\ i = \overline{1, n}\,.$$

Тоді

$$M\widehat{\mu}_n = \mu,\ D\widehat{\mu}_n = \frac{\sigma^2}{n}\,.$$

Тобто $\widehat{\mu}_n$ є незміщеною і слушною оцінкою μ. Далі

$$\|(I - \Pi_1)\vec{\xi}\|^2 = \|(\vec{\xi} - \vec{m}) - \Pi_1(\vec{\xi} - \vec{m})\|^2 = \|\vec{\xi} - \vec{m}\|^2 -$$
$$-\|\Pi_1(\vec{\xi} - \vec{m})\|^2 = \sum_{i=1}^n (\xi_i - \mu)^2 - n\left(\frac{1}{n}\sum_{i=1}^n (\xi_i - \mu)\right)^2,$$

де $\vec{m} = (\mu, \ldots, \mu)$. Тоді

$$M\|(I - \Pi_1)\vec{\xi}\|^2 = n\sigma^2 - \sigma^2 = (n-1)\sigma^2\,,$$

і відповідно до (8.34) маємо

$$M\widehat{\sigma}_n^2 = \frac{1}{n-1}(n-1)\sigma^2 = \sigma^2\,.$$

Тобто, і без припущення про нормальність розподілу, статистика $\widehat{\sigma}_n^2$ є незміщеною оцінкою σ_n^2. Але без додаткових припущень про розподіл нічого не можна сказати про слушність оцінки $\widehat{\sigma}_n^2$. Що стосується $\widetilde{\sigma}_n^2$, то у цьому випадку вона втрачає свої переваги перед $\widehat{\sigma}_n^2$.

Таким чином, важливими для нас є такі властивості оцінок:

1) *Незміщенність*: $M_{\vec{\theta}}t(\vec{\xi}) = \tau(\vec{\theta})$;

2) *Мінімальність відхилення*: $M_{\vec{\theta}}\left(t(\vec{\xi}) - \tau(\vec{\theta})\right)^2 \sim \min$. Зокрема, для незміщених оцінок це мінімальність дисперсії;

3) *Слушність*: $t_n(\vec{\xi}) \xrightarrow{p} \tau(\vec{\theta}),\ n \to \infty$.

8.7. Ефективні оцінки

При деяких припущеннях щодо функцій розподілу $F(x, \vec{\theta})$ може бути отримана зручна апріорна оцінка для якості оцінювання, яка відома як *нерівність Рао-Крамера.* Нехай $\vec{\xi}=(\xi_1, .., \xi_n)$ - випадкова вибірка об'ємом n із розподілу з густиною $p(x, \vec{\theta})$, $x \in R_1$, $\vec{\theta} \in \Theta \subset R_k$. Оскільки $\xi_1, \ldots, \xi_n$ незалежні у сукупності, то густина спільного розподілу дорівнює ($\vec{x} = (x_1, \ldots, x_n)$)

$$L(\vec{x}, \vec{\theta}) = p(x_1, \vec{\theta}) \ldots p(x_n, \vec{\theta})\,.$$

У статистичних задачах $L(\vec{x}, \vec{\theta})$ розглядається як функція двох аргументів $\vec{x} \in R_n$, $\vec{\theta} \in \Theta$, і називається *функцією правдоподібності.*

Якщо вибірку $\xi_1, \ldots, \xi_n$ утворюють дискретні випадкові величини, то функція правдоподібності визначається як добуток ймовірностей

$$L(\vec{x}, \vec{\theta}) = P(x_1, \vec{\theta}) \ldots P(x_n, \vec{\theta})\,.$$

Теорема 22. (Рао-Крамера) *Нехай $\Theta = R_1$; нехай статистика $t(\vec{\xi})$ є незміщеною оцінкою $\tau(\theta): \; Mt(\vec{\xi}) = \tau(\theta)$; нехай функції $L(\vec{x}, \theta)$ і $\tau(\theta)$ диференційовані по θ; нехай множина усіх тих $\vec{x} \in R_n$, для яких $L(\vec{x}, \theta) > 0$, не залежить від $\theta \in \Theta$ і нехай*

$$\frac{d}{d\theta} \int L(\vec{x}, \vec{\theta}) d\vec{x} = \int \frac{\partial L(\vec{x}, \vec{\theta})}{\partial \theta} d\vec{x}$$

та

$$\frac{d}{d\theta} \int t(\vec{x}) L(\vec{x}, \vec{\theta}) d\vec{x} = \int t(\vec{x}) \frac{\partial L(\vec{x}, \vec{\theta})}{\partial \theta} d\vec{x}\,.$$

Тоді

$$D_\theta t(\vec{\xi}) \geq \frac{\left[\dfrac{d\tau(\theta)}{d\theta}\right]^2}{M_\theta \left[\dfrac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta}\right]^2}\,, \tag{8.35}$$

причому знак рівності стоїть тоді і тільки тоді, коли рівність

$$\frac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta} = a(\theta) \left[t(\vec{\xi}) - \tau(\theta)\right] \tag{8.36}$$

виконується з ймовірністю одиниця для деякої функції $a(\theta)$, $\theta \in \Theta$.

Доведення не приводимо.

Незміщена оцінка $\tau(\theta): \; Mt(\vec{\xi}) = \tau(\theta)$ називається *ефективною*, якщо

$$D_\theta t(\vec{\xi}) = \frac{\left[\dfrac{d\tau(\theta)}{d\theta}\right]^2}{M_\theta \left[\dfrac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta}\right]^2}\,. \tag{8.37}$$

Відповідно до (8.36) має місце співвідношення

$$M_\theta \left(\frac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta} \right)^2 = M_\theta \left(a(\theta) \left[t(\vec{\xi}) - \tau(\theta) \right] \right)^2 = a^2(\theta) D_\theta t(\vec{\xi}) \,.$$

Підставляючи його у (8.37) отримаємо

$$D_\theta t(\vec{\xi}) = \left| \frac{1}{a(\theta)} \frac{d\tau(\theta)}{d\theta} \right| . \tag{8.38}$$

Приклад 47. *Знайти ефективну оцінку математичного сподівання для вибірки* $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ *із розподілу* $N(\theta, \sigma^2)$ *із відомою дисперсією* σ^2.

◁ У цьому випадку функція правдоподібності має вигляд

$$L(\vec{x}, \theta) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left[-\frac{1}{2\sigma^2} \sum_i (x_i - \theta)^2 \right] , \tag{8.39}$$

і, як наслідок,

$$\frac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left[-\frac{1}{2\sigma^2} \sum_i (x_i - \theta)^2 \right] = \frac{n}{\sigma^2} \left(\frac{1}{n} \sum_i x_i - \theta \right) .$$

Тоді, відповідно до (8.36), статистика $t(\vec{\xi}) = \dfrac{1}{n} \sum\limits_i \xi_i$ є ефективною оцінкою свого математичного сподівання θ. Відповідно до (8.38) дисперсія цієї оцінки дорівнює

$$D_\theta t(\vec{\xi}) = \left| \frac{1}{\left(\frac{n}{\sigma^2} \right)} \right| = \frac{\sigma^2}{n} \,.$$

▷

Приклад 48. *Знайти ефективну оцінку дисперсії для вибірки* $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ *із розподілу* $N(\mu, \theta^2)$ *з відомим математичним сподіванням* μ.

◁ У цьому випадку функція правдоподібності має вигляд

$$L(\vec{x}, \theta) = \left(\frac{1}{\sqrt{2\pi}\theta} \right)^n \exp \left[-\frac{1}{2\theta^2} \sum_i (x_i - \mu)^2 \right] ,$$

і, як наслідок,

$$\frac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^3} \sum_i (x_i - \mu)^2 = \frac{n}{\theta^3} \left[\frac{1}{n} \sum_i (x_i - \mu)^2 - \theta^2 \right] .$$

Тоді, відповідно до (8.36), статистика $t(\vec{\xi}) = \dfrac{1}{n}\sum_i (\xi_i - \mu)^2$ є ефективною оцінкою дисперсії θ^2. Відповідно до (8.38) дисперсія цієї оцінки дорівнює

$$D_\theta t(\vec{\xi}) = \left| \frac{2\theta}{\left(\frac{n}{\theta^3}\right)} \right| = \frac{2\theta^4}{n}\,.$$

Теорема (8.35) залишається вірною, якщо густину $p(\vec{x}, \theta)$ замінити на ймовірності, а інтегрування на підсумовування. $\rhd$

Приклад 49. *Знайти ефективну оцінку параметра θ для вибірки $\vec{\xi}=(\xi_1, \ldots, \xi_n)$ із розподілу Пуассона.*

$\lhd$ У цьому випадку функція правдоподібності має вигляд

$$L(\vec{x}, \theta) = \prod_i \frac{\theta^{x_i}}{x_i!} e^{-\theta}\,,$$

і, як наслідок,

$$\frac{\partial \ln L(\vec{\xi}, \theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \sum_i (-\theta + x_i \ln \theta - \ln x_i!) = \frac{n}{\theta}\left(\frac{1}{n}\sum_i x_i - \theta\right).$$

Тоді, відповідно до (8.36), статистика $t(\vec{\xi}) = \dfrac{1}{n}\sum_i \xi_i$ є ефективною оцінкою параметра θ. Відповідно до (8.38) дисперсія цієї оцінки дорівнює

$$D_\theta t(\vec{\xi}) = \left| \frac{1}{n/\theta} \right| = \frac{\theta}{n}\,.$$

$\rhd$

8.8. Оцінки максимальної правдоподібності

Нехай $L(\vec{x}, \vec{\theta})$, $\vec{\theta} \in \Theta$, $\vec{x} \in R_n$ - функція правдоподібності. *Оцінкою максимальної правдоподібності* називається статистика $\widehat{\vec{\theta}} = \widehat{\vec{\theta}}(\vec{\xi})$, яка задовольняє умові: $L(\vec{x}, \widehat{\vec{\theta}}) \geq L(\vec{x}, \vec{\theta})$ для усіх $\vec{\theta} \in \Theta$, або, що те ж саме,

$$L(\vec{\xi}, \widehat{\vec{\theta}}) = \max_{\vec{\theta} \in \Theta} L(\vec{\xi}, \vec{\theta})\,.$$

Нехай Θ - підмножина k - вимірного евклідового простору R_k, функція правдоподібності $L(\vec{x}, \vec{\theta})$ диференційована по $\vec{\theta} \in \Theta$ і досягає максимуму по $\vec{\theta}$ у внутрішній точці Θ для кожного $\vec{x} \in R_n$. У цьому випадку оцінка максимальної правдоподібності задовольняє рівнянням

$$\left.\frac{\partial \ln L(\vec{x}, \vec{\theta})}{\partial \theta_i}\right|_{\vec{\theta} = \widehat{\vec{\theta}}(\vec{x})} = 0\,, \ \ i = \overline{1, k}\,. \tag{8.40}$$

Рівняння (8.40) є необхідними умовами максимуму і називаються *рівняннями максимальної правдоподібності.*

Якщо $\vec{x} = (x_1, \ldots, x_n)$ - спостережене значення вибірки $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ із дискретного розподілу з параметрами $\vec{\theta}$, то $L(\vec{x}, \vec{\theta}) = P\{\xi_1 = x_1, \ldots, \xi_n = x_n | \vec{\theta}\}$ ймовірність того, що $\vec{x} = \vec{\xi}$. За оцінку максимальної правдоподібності $\widehat{\vec{\theta}}(\vec{x})$ приймається те значення параметра $\vec{\theta}$, при якому ймовірність отримати спостережене значення $\vec{x}$ вибірки $\vec{\theta}$ приймає максимальне значення. Це пояснює зміст принципу максимальної правдоподібності: як значення невідомого параметра пропонується приймати те, при якому ймовірність спостереженої реалізації вибірки приймає максимальне значення.

Приклад 50. *Знайти оцінку максимальної правдоподібності для $\vec{\theta} = (\mu, \sigma^2)$ на підставі вибірки $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ із $N(\mu, \sigma^2)$.*

$\triangleleft$ Оскільки

$$L(\vec{x}, \theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2\right],$$

то (8.40) приймає вигляд

$$\frac{\partial \ln L(\vec{x}, \vec{\theta})}{\partial \mu} = \frac{1}{\sigma^2}\sum_i (x_i - \mu) = 0\,,$$

$$\frac{\partial \ln L(\vec{x}, \vec{\theta})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_i (x_i - \mu)^2 = 0\,.$$

Звідси випливає оцінка максимальної правдоподібності

$$\widehat{\vec{\theta}} = (\widetilde{\mu}, \widetilde{\sigma}^2)\,, \quad \widetilde{\mu} = \frac{1}{n}\sum_i x_i\,, \quad \widetilde{\sigma}^2 = \frac{1}{n}\sum_i (x_i - \widetilde{\mu})^2\,.$$

$\triangleright$

8.9. Достатні статистики

Звернемо увагу на те, що при побудові точкових оцінок найчастіше не використовується вся інформація, що міститься у виборці. Наприклад, для оцінки математичного сподівання за вибіркою достатньо знати лише значення суми $\xi_1 + \ldots + \xi_n$, а не кожної випадкової величини ξ_i окремо. Аналогічно, при відомому μ для оцінки дисперсії достатньо знати суму $\sum_i (\xi_i - \mu)^2$ тощо.

Виявляється, що в багатьох випадках для оцінювання параметра $\vec{\theta}$ функції розподілу $F(x, \vec{\theta})$ за вибіркою $\xi_1, \ldots, \xi_n$ можна вказати статистику

$$\vec{T}(\vec{\xi}) = \Big(T_1(\vec{\xi}), \ldots, T_m(\vec{\xi})\Big)\,, \; m < n\,, \tag{8.41}$$

яка у деякому розумінні містить усю інформацію про параметр $\vec{\theta}$.

Означення. Статистика $\vec{T}(\vec{\xi})$ (8.41) називається *достатньою* для родини густин (ймовірностей) $L(\vec{x}, \vec{\theta})$, $\vec{\theta} \in \Theta$, $\vec{x} \in R_n$, якщо умовний розподіл випадкового вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ за умови $\vec{T}(\vec{\xi}) = \vec{t}$ не залежить від $\vec{\theta}$. На практиці зручною є така необхідна і достатня умова достатності статистики $\vec{T}(\vec{\xi})$.

Теорема 23. (Про факторизацію) $\vec{T}(\vec{\xi})$ *є достатньою статистикою тоді і тільки тоді, коли функція правдоподібності може бути представлена у вигляді:*

$$L(\vec{x}, \vec{\theta}) = g\left(\vec{T}(\vec{x}), \vec{\theta}\right) \cdot h(\vec{x}),\ \vec{x} \in R_n,\ \vec{\theta} \in \Theta\,.$$

Доведення не приводимо.

Приклад 51. *Знайти достатню статистику для нормального розподілу.*

$\lhd$ У випадку розподілу $N(\theta, \sigma^2)$ функція правдоподібності має вигляд (8.39)

$$\begin{aligned}
L(\vec{x}, \theta) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_i (x_i - \theta)^2\right] = \\
&= \left\{\exp\left[\frac{\theta}{\sigma^2}\sum_i x_i - \frac{n\theta^2}{2\sigma^2}\right]\right\} \cdot \left\{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_i x_i^2\right]\right\} = \\
&= g\left(\vec{T}(\vec{x}), \vec{\theta}\right) \cdot h(\vec{x})\,.
\end{aligned}$$

Таким чином, у нашому випадку, $\vec{T}(\vec{x}) = T(\vec{x}) = \sum_i x_i$ є достатньою статистикою. $\rhd$

8.10. Лінійна модель вимірів

Нехай в експерименті ми отримали випадковий вектор вимірів $\vec{\xi} = (\xi_1, \ldots, \xi_n)$

$$\xi_i = a_{i1}\alpha_1 + \ldots + a_{ik}\alpha_k + \nu_i, \quad k \le n\,, \tag{8.42}$$

де $\vec{\nu} = (\nu_1, \ldots, \nu_n)$ - випадковий вектор похибок. Коефіцієнти a_{ij} нам відомі і необхідно визначити вектор параметрів $\vec{\alpha} = (\alpha_1, \ldots, \alpha_k)$. Природно вважати, що $M\nu_i = 0$ (тобто систематичні похибки відсутні).

Нехай виміри (8.42) незалежні у сукупності і мають однакову точність, яка визначається дисперсією $M\nu_i^2 = \sigma^2$. Але тут дисперсія σ^2 невідома. Окрім цього, стовпчики прямокутної матриці $A = (a_{ij})$ будемо вважати лінійно незалежними. Задача полягає в тому, щоб за результатами вимірів $\vec{\xi}$ оцінити $\vec{\alpha}$ та σ^2. Це є так звана *обернена задача*, коли за даними вимірів необхідно визначити такі параметри об'єкта або явища, що безпосередньо не спостерігаються. В математичній статистиці ця задача називається *задачею аналізу лінійної регресії.*

Перепишемо (8.42) у вигляді

$$\vec{\xi} = A\vec{\alpha} + \vec{\nu}, \quad \vec{\alpha} \in R_k\,.$$

Лінійну незміщену оцінку параметра $\vec{\alpha}$ можна шукати у вигляді

$$\widehat{\vec{\alpha}} = B\vec{\xi}\,,$$

де B - невідома прямокутна матриця $(k \times n)$. Вимога незміщеності приводить до умови

$$M\widehat{\vec{\alpha}} = BM\vec{\xi} = B(AM\vec{\alpha} + M\vec{\nu}) = BA\vec{\alpha}\,.$$

Оскільки $\vec{\alpha}$ заздалегідь невідома, звідси випливає, що $M\widehat{\vec{\alpha}} = \vec{\alpha}$, і далі

$$(BA)_{ij} = \delta_{ij}\,, \qquad \text{(одинична матриця)}\,. \tag{8.43}$$

Обчислимо дисперсію компонент $\widehat{\alpha}_j$.

$$D\widehat{\alpha}_j = \sum_{i=1}^{n}(b_{ji})^2 D\xi_i = \sigma^2\sum_{i=1}^{n}(b_{ji})^2 = \sigma^2(BB^T)_{jj}\,. \tag{8.44}$$

Таким чином, задача визначення незміщеної оцінки мінімальної дисперсії приводиться до задачі на мінімум дисперсії (8.44) при обмеженнях (8.43). Для розв'язку цієї задачі використаємо метод невизначених множників Лагранжа. Нехай відповідна функція Лагранжа дорівнює

$$L = (BB^T)_{jj} - 2\sum_{l=1}^{k}\Lambda_{jl}\left[(BA)_{jl} - \delta_{jl}\right]\,,$$

де Λ_{jl} - множники Лагранжа. Маємо

$$\frac{\partial L}{\partial b_{ji}} = 2b_{ji} - 2\sum_{l=1}^{k}\Lambda_{jl}a_{il} = 0, \quad i = \overline{1,n}\,,$$

або у матричному вигляді $B = \Lambda A^T$. Використаємо (8.43)

$$(\Lambda A^T A)_{ij} = \sum_{l=1}^{k}\Lambda_{il}(A^T A)_{lj} = \delta_{ij}\,,$$

або $\Lambda A^T A = I$, $\Lambda = (A^T A)^{-1}$. І як результат $B = (A^T A)^{-1}A^T$, та лінійна незміщена оцінка мінімальної дисперсії має вигляд

$$\widehat{\vec{\alpha}} = (A^T A)^{-1}A^T\vec{\xi}. \tag{8.45}$$

Цікаво, що ця оцінка співпадає з результатом метода найменших квадратів, коли шукається мінімум по $\vec{\alpha}$ величини $\|\vec{\xi} - A\vec{\alpha}\|^2 \to \|\vec{\nu}\|^2$. Покажемо це. Нехай $\widetilde{\alpha}$ - оцінка вектора параметрів, яку можна знайти за допомогою метода найменших квадратів, тобто:

$$\min_{\widetilde{\vec{\alpha}}}\|\vec{\xi} - A\widetilde{\vec{\alpha}}\|^2\,. \tag{8.46}$$

Одну з компонент вектора, що стоїть під знаком норми у (8.46), можна записати у вигляді $\xi_i - \sum_{\omega=1}^{k} a_{i\omega}\widetilde{\alpha}_\omega$. Якщо визначити вектор

$$\vec{a}_\omega = \begin{pmatrix} a_{1\omega} \\ \vdots \\ a_{n\omega} \end{pmatrix}, \tag{8.47}$$

то (8.46) можна переписати

$$\min_{\widetilde{\vec{\alpha}}} \|\vec{\xi} - \sum_{\omega=1}^{k} \widetilde{\alpha}_\omega \vec{a}_\omega\|^2, \tag{8.48}$$

де множина $\vec{a}_\omega$ складає лінійну оболонку $L_a \subset R_n$.

Відповідно до (8.47) величину (8.48) можна розглядати як квадрат відстані $\rho(\vec{\xi}, L_a)$, тоді згідно (8.9) ця відстань дорівнює

$$\rho(\vec{\xi}, L_a) = \|\vec{\xi} - \Pi_a \vec{\xi}\|\,,$$

де Π_a- оператор проектування на L_a, а отже

$$\Pi_a \vec{\xi} = \sum_{\omega=1}^{k} \widetilde{\alpha}_\omega \vec{a}_\omega = A\widetilde{\vec{\alpha}}\,,$$

де $\widetilde{\alpha}_\omega$ мінімізує (8.48) , тобто

$$\frac{\partial}{\partial \widetilde{\alpha}_\omega} \|\vec{\xi} - \sum_{\gamma=1}^{k} \widetilde{\alpha}_\gamma \vec{a}_\gamma\|^2 = 2((\vec{\xi} - \sum_{\gamma=1}^{k} \widetilde{\alpha}_\gamma \vec{a}_\gamma), \vec{a}_\omega) = 0\,,$$

або

$$\sum_{\gamma=1}^{k} (\vec{a}_\gamma, \vec{a}_\omega) \widetilde{\alpha}_\gamma = (\vec{\xi}, \vec{a}_\omega)\,, \tag{8.49}$$

або

$$\sum_{\gamma=1}^{k} \sum_{i=1}^{n} a_{i\gamma} a_{i\omega} \widetilde{\alpha}_\gamma = \sum_{\gamma=1}^{k} \sum_{i=1}^{n} a^T_{\omega i} a_{i\gamma} \widetilde{\alpha}_\gamma = \sum_{\gamma=1}^{k} (A^T A)_{\omega\gamma} \widetilde{\alpha}_\gamma = [(A^T A)\widetilde{\vec{\alpha}}]_\omega\,.$$

Права частина (8.49) дорівнює

$$(\vec{\xi}, \vec{a}_\omega) = \sum_{i=1}^{n} \xi_i a_{i\omega} = \sum_{i=1}^{n} a^T_{\omega i} \xi_i = (A^T \vec{\xi})_\omega\,.$$

Тоді

$$(A^T A)\widetilde{\vec{\alpha}} = A^T \vec{\xi} \Rightarrow \widetilde{\vec{\alpha}} = (A^T A)^{-1} A^T \vec{\xi},$$

що співпадає з оцінкою (8.45) для $\widehat{\vec{\alpha}}$.

Розділ 9

Скінченні однорідні ланцюги Маркова

Нагадаємо, що послідовністю *n незалежних випробувань* ми називаємо ймовірнісний простір (Ω_n, F_n, P_n); елементами Ω_n є всі можливі послідовності $(\omega_{i_1}, \ldots, \omega_{i_n})$, де $\omega_{i_k} \in \Omega$, $\Omega = (\omega_1, \omega_2, \ldots)$ - множина елементарних подій в кожному випробуванні, F_n - σ-алгебра усіх підмножин Ω_n, при цьому незалежність випробувань зафіксована визначенням ймовірності P_n:

$$P_n(\omega_{i_1}, ..., \omega_{i_n}) = p_{i_1} \ldots p_{i_n} \,,$$

де $p_k = P(\{\omega_k\})$ - ймовірність елементарної події ω_k в одному випробуванні.

Припустимо, що в послідовності n випробувань ймовірність елементарної події в s-ому випробуванні *залежить* від елементарної події $(s-1)$-го випробування та *не залежить* від елементарних подій з номерами $s-2, s-3, \ldots, 1$. Позначимо цю умовну ймовірність

$$p_{ik} = P(\{\omega_k\} \mid \{\omega_i\}) \,, \quad 0 \le p_{ik} \le 1 \,.$$

У цьому позначенні використано властивість *однорідності послідовності випробувань*: ймовірність події ω_k в s-ому випробуванні при умові, що в $(s-1)$-ому випробуванні відбулася подія ω_i, не залежить від номера s.

Нехай заданий розподіл ймовірностей при першому випробуванні

$$P(\{\omega_j\}) = a_j \,; \quad j = 1, 2, \ldots; \quad a_j \ge 0; \quad \sum_{j=1}^{\infty} a_j = 1 \,.$$

Тоді ймовірність події $\{\omega_i \omega_j\}$ для першого і другого випробувань дорівнює

$$P(\{\omega_i \omega_j\}) = a_i p_{ij} \,,$$

і аналогічно для k випробувань

$$P(\{\omega_{i_1} \ldots \omega_{i_k}\}) = a_i p_{i_1 i_2} p_{i_2 i_3} .. p_{i_{k-1} i_k} \,, \; i_1, .., i_k = 1, 2, .. \tag{9.1}$$

Доведемо, що при $k = n$

$$\sum_{(\omega_{i_1} \ldots \omega_{i_n}) \in \Omega_n} P_n(\{\omega_{i_1} \ldots \omega_{i_n}\}) = \sum_{i_1=1,\ldots,i_n=1}^{\infty} a_{i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n} = 1 \,.$$

Дійсно

$$\sum_{(\omega_{i_1}..\omega_{i_n})\in\Omega_n} P_n(\{\omega_{i_1}..\omega_{i_n}\}) = \sum_{i_1=1,..,i_{n-1}=1}^{\infty} a_{i_1}p_{i_1i_2}..p_{i_{n-2}i_{n-1}} \sum_{i_n=1}^{\infty} p_{i_{n-1}i_n} =$$

$$= \sum_{i_1=1,\dots,i_{n-1}=1}^{\infty} a_{i_1}p_{i_1i_2}\dots p_{i_{n-2}i_{n-1}} = \dots = \sum_{i_1=1}^{\infty} a_{i_1} = 1\,.$$

Надалі для простоти покладемо, що простір Ω - скінченний, $\Omega = (\omega_1, \dots, \omega_N)$.

Скінченним однорідним ланцюгом Маркова, що складається із n випробувань, називається ймовірнісний простір (Ω_n, F_n, P_n), в якому Ω_n - множина усіх послідовностей $(\omega_{i_1}, .., \omega_{i_n})$, $1 \le i_1, \dots, i_n \le N$, F_n - алгебра усіх підмножин Ω_n і ймовірність P_n визначена для кожної елементарної події Ω_n за допомогою рівності (9.1) з $k = n$, де $a_i \ge 0$, $\sum_{i=1}^N a_i = 1$ - початковий розподіл; величини $p_{ij} \ge 0$, $\sum_{j=1}^N p_{ij} = 1$ - називаються *перехідними ймовірностями.*

Сукупність ймовірностей переходу p_{ij}, $i, j = \overline{1, N}$, утворює *матрицю переходу* Π_1

$$\Pi_1 = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & & \vdots \\ p_{N1} & \cdots & p_{NN} \end{pmatrix}.$$

За означенням $p_{ij} \ge 0$ і сума елементів кожного рядка дорівнює одиниці.

Розглянемо ймовірність переходу із стану ω_i, яке було реалізовано в s-ому випробуванні, в стан ω_j за n кроків, тобто в стан ω_j в $(s+n)$-ому випробуванні. Внаслідок однорідності ланцюга Маркова ця ймовірність залежить тільки від n (і не залежить від s). Позначимо її p_{ij}^n. Тоді p_{ik}^m є ймовірністю переходу $i \to k$ за m кроків і p_{kj}^{n-m} є ймовірністю переходу $k \to j$ за $(n-m)$ кроків. Очевидно, що відповідно до формули повної ймовірності

$$p_{ij}^n = \sum_{k=1}^N p_{ik}^m p_{kj}^{n-m}, \quad m = \overline{1, n-1}, \tag{9.2}$$

де враховувалась попарна несумісність подій $(\omega_1, \dots, \omega_n)$.

Позначимо Π_n матрицю, що складена із p_{ij}^n. Тоді (9.2) можна переписати

$$\Pi_n = \Pi_m \Pi_{n-m}, \quad m = \overline{1, n-1},$$

або

$$\Pi_n = \Pi_1\Pi_{n-1} = \Pi_1\Pi_1\Pi_{n-2} = \dots = (\Pi_1)^n \tag{9.3}$$

- *ймовірність переходу за n кроків.*

Розглянемо окремий випадок: *незалежні випробування.* Тоді

$$\Pi_1 = \begin{pmatrix} a_1 & a_2 & \dots & a_N \\ a_1 & a_2 & \dots & a_N \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \dots & a_N \end{pmatrix}. \tag{9.4}$$

Далі маємо

$$
\begin{aligned}
\Pi_1\Pi_1 &= \begin{pmatrix} a_1 & a_2 & \dots & a_N \\ a_1 & a_2 & \dots & a_N \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \dots & a_N \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \dots & a_N \\ a_1 & a_2 & \dots & a_N \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \dots & a_N \end{pmatrix} = \\
&= \begin{pmatrix} a_1(a_1 + a_2 \dots a_n) & a_2(a_1 + a_2 \dots a_n) & \dots \\ a_1(a_1 + a_2 \dots a_n) & a_2(a_1 + a_2 \dots a_n) & \dots \\ \vdots & \vdots & \vdots \\ a_1(a_1 + a_2 \dots a_n) & a_2(a_1 + a_2 \dots a_n) & \dots \end{pmatrix} = \\
&= \begin{pmatrix} a_1 & a_2 & \dots & a_N \\ a_1 & a_2 & \dots & a_N \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \dots & a_N \end{pmatrix} = \Pi_1 \,,
\end{aligned}
\tag{9.5}
$$

тобто

$$\Pi_1 = \Pi_n = \Pi_1^n \,.$$

Іншими словами, перехід із i-го в j-ий стан не залежить від i:

$$p_{ij}^n = p_j \,.$$

Інтуїтивно можна очікувати, що і у випадку довільного ланцюга Маркова при переходах в n кроків вплив початкового розподілу повинен зменшуватися у тому розумінні, що

$$p_{ij}^n \to p_j \quad \text{при} \quad n \to \infty$$

незалежно від i. Така гранична матриця переходів має вигляд матриці (9.4) і задовольняє умову

$$\Pi_1^k = \Pi_1, \quad k = 1, 2, \dots \tag{9.6}$$

Умова (9.6) називається *ергодичністю.*

Нехай $a_1 \dots a_N$ - початковий розподіл ймовірностей. Визначимо p_j^k - *абсолютну ймовірність* (ймовірність того, що в k-ому випробуванні реалізується ω_j). Тоді

$$
\begin{aligned}
&p_j^2 = \textstyle\sum_{k=1}^N a_k p_{kj} \,, \\
&p_j^3 = \textstyle\sum_{k=1}^N p_k^2 p_{kj} = \sum_{k=1}^N \left(\sum_{l=1}^N a_l p_{lk} \right) p_{kj} = \\
&\quad = \textstyle\sum_{l=1}^N a_l \left(\sum_{k=1}^N p_{lk} p_{kj} \right) = \sum_{l=1}^N a_l p_{lj}^2 \,, \\
&\dots \\
&p_j^n = \textstyle\sum_{k=1}^N p_k^{n-1} p_{kj} = \sum_{l=1}^N a_l p_{lj}^{n-1} \,.
\end{aligned}
\tag{9.7}
$$

Якщо існують границі при $n \to \infty$

$$\lim_{n\to\infty} p_{kj}^n = p_j, \quad k = \overline{1, N},$$

то $$\lim_{n\to\infty} p_j^n = \sum_{l=1}^{N} a_l \lim_{n\to\infty} p_{lj}^{n-1} = \sum_{l=1}^{N} a_l p_j = p_j.$$

Тобто, якщо існує *ергодичність*, то існує граничний *фінальний* розподіл ймовірностей $p_1 \dots p_N$, що не залежить від початкового розподілу $a_1 \dots a_N$.

Таким чином, кінцевий розподіл ймовірностей $p_1 \dots p_N$ задовольняє системі рівнянь

$$p_j = \lim_{n\to\infty} p_j^n = \sum_{l=1}^{N} \lim_{n\to\infty} p_l^{n-1} p_{lj} = \sum_{l=1}^{N} p_l p_{lj} \tag{9.8}$$

з очевидними додатковими умовами

$$p_j \geq 0\,, \quad \sum_{j=1}^{N} p_j = 1\,.$$

При порівнянні (9.7) і (9.8) видно, що, якщо як початковий розподіл $a_1, \dots, a_N$ обрати фінальний розподіл $p_1, \dots, p_N$, то розподіл станів $\omega_1, \dots, \omega_N$ не буде змінюватись від випробування до випробування. Тобто, *фінальний розподіл стаціонарний.*

Справедлива наступна рівність, що є наслідком ергодичності. Нехай в кожному стані до переходу в наступний стан система знаходиться τ секунд. Нехай A - деяка множина станів, T_A - час, на протязі якого система знаходиться в станах із множини A, $T = n\tau$ - загальний час функціонування системи. Тоді, у випадку ергодичності

$$\lim_{n\tau = T\to\infty} \left(\frac{T_A}{T}\right) = \sum_{\omega_i \in A} p_i.$$

Приклад 52. *Нехай є ланцюг Маркова з двома станами ω_1 та ω_2, ймовірностями переходу $p_{11} = p_{22} = p$, $p_{12} = p_{21} = q$, $(0 \leq p \leq 1,\ p + q = 1)$ і початковими ймовірностями $p(\omega_1) = a_1$, $p(\omega_2) = a_2$, $a_1 + a_2 = 1$. Знайти матрицю переходу через n випробувань, абсолютні ймовірності p_i^n і фінальні ймовірності p_i, $i = 1, 2$.*

$\lhd$ В силу (9.3) маємо $\Pi_n = \Pi_1^n$. Нехай $\Pi_1 = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$ - матриця лінійного оператора A в евклідовому просторі R_2 із базисом $\vec{f_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ і $\vec{f_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Знайдемо власні вектори і власні числа оператора A:

$$A\vec{e} = \lambda\vec{e},$$

де $\vec{e} = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$, або $\begin{vmatrix} p-\lambda & q \\ q & p-\lambda \end{vmatrix} = 0\,,$

$$(p-\lambda)^2 = q^2\,, \qquad \left.\begin{array}{ll} p - \lambda_1 = -q\,, & \lambda_1 = p + q = 1\,, \\ p - \lambda_2 = +q\,, & \lambda_2 = p - q = c\,. \end{array}\right\}$$

Тоді, для $\lambda = \lambda_1 = 1$, маємо $\left.\begin{array}{rcl} p\xi + q\eta & = & \xi \\ q\xi + p\eta & = & \eta \end{array}\right\}$. Якщо $\eta = 1$, то $\xi = \dfrac{q}{1-p} = \dfrac{q}{q} = 1$, тобто

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Для $\lambda = \lambda_2 = c$, маємо $\left.\begin{array}{rcl} p\xi + q\eta & = & c\xi \\ q\xi + p\eta & = & c\eta \end{array}\right\}$. Якщо $\eta = -1$, то $\xi = \dfrac{(c-p)(-1)}{q} = \dfrac{(p-q-p)(-1)}{q} = 1$, тобто

$$\vec{e}_2 = \begin{pmatrix} +1 \\ -1 \end{pmatrix}.$$

В базисі $\{\vec{e}_1, \vec{e}_2\}$ матриця $\hat{\Pi}_1$ діагональна

$$\hat{\Pi}_1 = \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix} \quad \text{та} \quad \hat{\Pi}_n = \hat{\Pi}_1^n = \begin{pmatrix} 1 & 0 \\ 0 & c^n \end{pmatrix}.$$

Якщо B - матриця переходу від базису $\{\vec{e}_1, \vec{e}_2\}$ до базису $\{\vec{f}_1, \vec{f}_2\}$, тобто для довільного вектора $\vec{X}$ має місце зв'язок між його координатами в різних базисах

$$\begin{array}{rcccl} \vec{X} & = & B\vec{X}' & = & x_1\vec{e}_1 + x_2\vec{e}_2\,, \\ \vec{X}' & = & B^{-1}\vec{X} & = & x_1'\vec{f}_1 + x_2'\vec{f}_2\,, \end{array}$$

то справедливі співвідношення

$$\begin{array}{rcl} \vec{Y} & = & \hat{\Pi}_n\vec{X}\,, \\ \vec{Y}' & = & B^{-1}\left(\hat{\Pi}_n\vec{X}\right) \; = \; B^{-1}\hat{\Pi}_n B B^{-1}\vec{X} \; = \; \Pi_n\vec{X}'\,, \\ \Pi_n & = & B^{-1}\hat{\Pi}_n B\,. \end{array}$$

Знайдемо компоненти матриці B. Якщо $B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, то ці компоненти можна знайти таким чином:

$$\vec{f}_1 = B\vec{e}_1\,, \quad \vec{f}_2 = B\vec{e}_2\,,$$

або

$$\left.\begin{array}{rclcrcl} 1 & = & \alpha + \beta\,, & & 0 & = & \alpha - \beta \\ 0 & = & \gamma + \delta\,, & & 1 & = & \gamma - \delta \end{array}\right\},$$

тобто $\alpha = \beta = \dfrac{1}{2}$, $\quad \gamma = -\delta = \dfrac{1}{2}$. Звідси

$$B = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{та} \quad B^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Тут ми ще раз отримали підтвердження відомого результату, а саме, перехід до базису, що побудований на власних векторах симетричної матриці, здійснюється за допомогою матриці (в даному випадку B), стовпчики якої пропорційні елементам власних векторів цієї симетричної матриці (в даному випадку векторів $\vec{e}_1$ та $\vec{e}_2$). Тоді

$$\Pi_n = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & c^n \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+c^n & 1-c^n \\ 1-c^n & 1+c^n \end{pmatrix}.$$

Таким чином

$$\left.\begin{aligned} p_{11}^n = (\Pi_n)_{11} = p_{22}^n = (\Pi_n)_{22} = \frac{1}{2}\left[1+(p-q)^n\right] \\ p_{12}^n = (\Pi_n)_{12} = p_{21}^n = (\Pi_n)_{21} = \frac{1}{2}\left[1-(p-q)^n\right] \end{aligned}\right\} n = 1, 2, \ldots$$

Тоді, внаслідок (9.7), маємо

$$\begin{aligned} p_1^{n+1} &= a_1 p_{11}^n + a_2 p_{21}^n = a_1 \cdot \frac{1}{2}\left[1+(p-q)^n\right] + a_2 \cdot \frac{1}{2}\left[1-(p-q)^n\right] = \\ &= \frac{1}{2}(a_1+a_2) + \frac{1}{2}(a_1-a_2)(p-q)^n = \frac{1}{2} + \frac{1}{2}(a_1-a_2)(p-q)^n, \end{aligned}$$

і аналогічно

$$p_2^{n+1} = a_1 p_{12}^n + a_2 p_{22}^n = \frac{1}{2} - \frac{1}{2}(a_1-a_2)(p-q)^n.$$

Звідки

$$\left.\begin{array}{ccccc} \lim\limits_{n\to\infty} p_{i1}^n & = & \lim\limits_{n\to\infty} p_{i2}^n & = & \dfrac{1}{2} \\ \lim\limits_{n\to\infty} p_1^n & = & \lim\limits_{n\to\infty} p_2^n & = & \dfrac{1}{2} \end{array}\right\}.$$

Отже, для приклада 52. *фінальний розподіл стаціонарний.* $\triangleright$

Розділ 10

Випадкові процеси

Означення. Нехай (Ω, F, P) - ймовірнісний простір, T - деяка числова множина. Дійсна функція $\xi(t) = f(t,\omega)$, що визначена при $t \in T, \omega \in \Omega$, називається *випадковим процесом (випадковою функцією)*, якщо при кожному $t \in T$ значення $f(t,\omega)$ як функція $\omega \in \Omega$ є випадковою величиною.

Якщо фіксовано $\omega = \omega_0$, то $f(t,\omega_0)$ - функція від $t \in T$ називається *реалізацією* випадкового процесу $\xi(t)$, або *вибірковою функцією.* Якщо фіксовано $t = t_0$, то випадкова величина $\xi(t_0)$ називається *перерізом* випадкового процесу в точці $t = t_0$.

В кожному перерізі розподіл ймовірностей випадкового процесу задається одновимірною функцією розподілу $F(t,x) = = P\{\xi(t) < x\}$. Більш повний опис дає k - вимірна функція розподілу випадкового вектора $\vec{\xi} = (\xi(t_1), \ldots, \xi(t_k))$:

$$F(t_1, x_1; \ldots; t_k, x_k) = P\{\xi(t_1) < x_1, \ldots, \xi(t_k) < x_k\}.$$

Розглянуті вище послідовності незалежних випадкових величин $\xi_1, \ldots, \xi_n$ або ланцюги Маркова уявляють собою приклади випадкових процесів.

10.1. Процес Пуассона

Це є приклад процесу із *незалежними приростами.* Розглянемо деяку подію A, яка може відбуватися у випадкові моменти часу, і нехай $\xi(t)$ - кількість настання події A у проміжку часу довжиною t. Нехай виконано умови:

1) Випадкові величини $\xi(t_j),\ j = 1, 2, \ldots$ для проміжків часу, що не перетинаються, незалежні у сукупності.

2) Для довільного проміжку часу ймовірність настання події A у цьому проміжку залежить тільки від довжини цього проміжку (і не залежить від того, де на осі часу він розташований) - властивість *однорідності* часу.

3) При $t \to 0$ має місце умова

$$P\{\xi(t) = 1\} = \lambda t + o(t), \quad \lambda > 0, \quad P\{\xi(t) > 1\} = +o(t).$$

Умови 1)-3) виконуються для розпаду радіоактивної речовини, відмови радіоустаткування, викликів на телефонній станції і т.д.

Теорема 24. *Нехай виконуються умови 1)-3). Розподіл випадкового вектора* $\vec{\xi} = (\xi(t_1),\ldots,\xi(t_j))$ *для проміжків часу, що не перетинаються, є пуассонівським:*

$$P\{\xi(t_1)=k_1,\ldots,\xi(t_j)=k_j\}=\frac{(\lambda t_1)^{k_1}}{k_1!}\cdots\frac{(\lambda t_j)^{k_j}}{k_j!}e^{-\lambda(t_1+\ldots+t_j)}\,. \tag{10.1}$$

$\lhd$ Візьмемо довільний інтервал $t_l = t$ і розіб'ємо його на n інтервалів $\Delta_i = \Delta = \dfrac{t}{n}$, що не перетинаються, і потім спрямуємо $n \to \infty$. Подію $\{\xi(t) = k\}$ представимо у вигляді

$$\{\xi(t) = k\} = A_1 + A_2\,, \tag{10.2}$$

де

$$A_1=\bigcup_{i_1,\ldots,i_k}\{\xi(\Delta_{i_1})=1,\ldots,\xi(\Delta_{i_k})=1,\xi(\Delta_{i_{k+1}})=0,\ldots,\xi(\Delta_{i_n})=0\}\ . \tag{10.3}$$

Тобто, A_1 - це така подія, коли в проміжку Δ_i відбувається не більше однієї події. Подія A_2 - це сума всіх інших подій, тобто принаймні в одному Δ_i подія A відбувається не менше двох разів:

$$A_2 \subset \bigcup_{i=1}^{n}\{\xi(\Delta_i) > 1\}\ .$$

Внаслідок властивості 2) має місце оцінка

$$\begin{aligned} P(A_2) &\le P\left(\bigcup_{i=1}^{n}\{\xi(\Delta_i) > 1\}\right) \le \sum_{i=1}^{n} P\{\xi(\Delta_i) > 1\} = \\ &= n \cdot o(\frac{t}{n}) = o(1)\,, \quad n \to \infty\,. \end{aligned} \tag{10.4}$$

Нехай

$$p_0 = P\{\xi(\Delta) = 0\}\ , \quad p_1 = P\{\xi(\Delta) = 1\}\ .$$

Відповідно до визначення (10.3) події, що входять в A_1, попарно несумісні. Тоді із 1) випливає, що

$$P(A_1) = C_n^k p_1^k p_0^{n-k}\,. \tag{10.5}$$

Як наслідок, із (10.2), (10.4) і (10.5) можна записати

$$P\{\xi(t) = k\} = C_n^k p_1^k p_0^{n-k} + o(1)\,, \quad n \to \infty\,.$$

З умови 3) випливає, що

$$\begin{aligned} &p_0=1-P\{\xi(\Delta)\ge 1\} - P\{\xi(\Delta)=1\} - P\{\xi(\Delta)>1\} = 1-\lambda\Delta + o(\Delta),\\ &p_1 = P\{\xi(\Delta) = 1\} = \lambda\Delta + o(\Delta)\,, \quad \Delta = \tfrac{t}{n}\,. \end{aligned}$$

Повторюючи доведення теореми Пуассона (формула (3.9)) отримаємо

$$P\{\xi(\Delta)=k\}=C_n^k(\lambda\Delta)^k(1-\lambda\Delta)^{n-k}\longrightarrow\frac{(\lambda t)^k}{k!}e^{-\lambda t}\,.$$

Тоді внаслідок властивості 1) має місце рівність (10.1). ▷

Аргументом t випадкового процесу $\xi(t)$ є *довжина* проміжку часу. Зважаючи на однорідність процесу як проміжок можна взяти інтервал $(0,t)$. Позначимо $\eta(t)=\xi(t)$, де аргументом процесу $\eta(t)$ буде поточний момент часу t, тобто наприклад, $\eta(t)$ відображає кількість α–частинок, що були зареєстровані до моменту часу t і таке інше.

Означення. Випадковий процес $\eta(t)$, $0\le t<\infty$ називається *процесом Пуассона*, якщо:

а) для довільної послідовності $0\le t_1<t_2<\ldots<t_n$ випадкові величини $\eta(t_2)-\eta(t_1),\ldots,\eta(t_n)-\eta(t_{n-1})$ незалежні у сукупності (такі процеси називаються *процесами з незалежними приростами*);

б) випадкова величина $\eta(t)-\eta(s)$ має розподіл Пуассона з параметром $\lambda(t-s)$

$$P\{\eta(t)-\eta(s)=k\}=\frac{[\lambda(t-s)]^k}{k!}e^{-\lambda(t-s)}\,. \tag{10.6}$$

Якщо додатково вимагати, щоб $\eta(0)=0$, то відповідна випадкова величина як функція часу буде мати вигляд, що зображений на Рис. 10.1. Тут точки $t_1,t_2,\ldots$ відповідають моментам часу, коли відбувається випадкова подія.

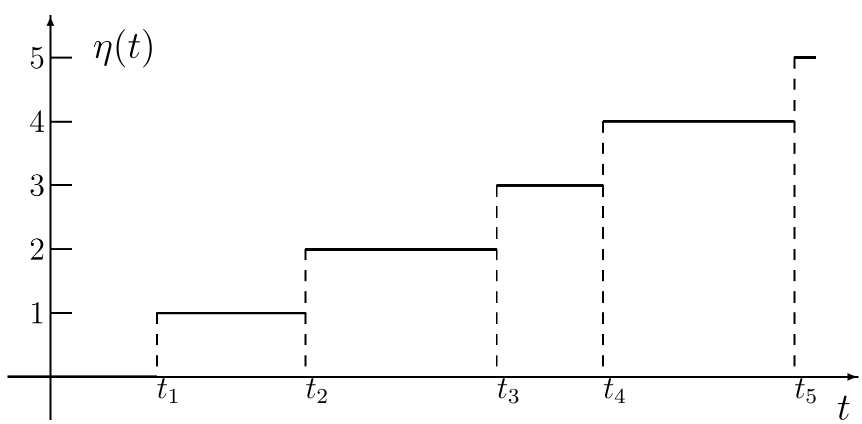

Рис. 10.1.

Нехай τ - час очікування першої події в пуассонівському потоці подій. Знайдемо функцію розподілу випадкової величини τ. Відповідно до (10.6) маємо

$$P\{\tau\ge t\}=P\{\eta(t)-\eta(0)=0\}=e^{-\lambda t}\,.$$

Тоді функція розподілу

$$F_\tau(t)=P\{\tau<t\}=1-e^{-\lambda t}\,,$$

і густина розподілу

$$p_\tau(t) = \frac{d}{dt}F_\tau(t) = \lambda e^{-\lambda t}, \quad 0 \le t < \infty \, .$$

Якщо зсунути початковий час, то

$$P\{\tau \ge t\} = P\{\eta(t+t_1) - \eta(t_1) = 0\} = e^{-\lambda t}\,,$$

і внаслідок цього час очікування *чергової* події не зменшується від того, що ця подія вже очікувалась і до моменту t_1. Дійсно, оскільки $\{\tau - t_1 \ge t\} \bigcap \{\tau \ge t_1\} = \{\tau - t_1 \ge t\}$, то

$$P\{\tau - t_1 \ge t | \tau \ge t_1\} = \frac{P\left\{\{\tau - t_1 \ge t\} \bigcap \{\tau \ge t_1\}\right\}}{P\{\tau \ge t_1\}} = \frac{P\{\tau - t_1 \ge t\}}{P\{\tau \ge t_1\}} =$$
$$= \frac{e^{-\lambda(t+t_1)}}{e^{-\lambda t_1}} = e^{-\lambda t} = P\{\tau \ge t\}\,.$$

Обчислимо $M\tau$

$$M\tau = \int_0^\infty t p_\tau(t) dt = \int_0^\infty t\lambda e^{-\lambda t} dt = \frac{1}{\lambda}\,,$$

тобто λ - середня кількість подій в одиницю часу.

10.2. Процес Вінера

Означення. Процес $\xi(t)$, $0 \le t < \infty$ називається *процесом Вінера*, якщо виконуються такі умови:
1) для довільних $0 \le t_0 < t_1 < \ldots < t_n$ випадкові величини $\eta_1 = \xi(t_1) - \xi(t_0), \ldots, \eta_n = \xi(t_n) - \xi(t_{n-1})$ незалежні у сукупності;
2) випадкова величина $\xi(t) - \xi(s)$ $(0 \le s < t$ - довільні$)$ має нормальний розподіл $N(0, t-s)$;
3) в початковий момент часу $\xi(0) = 0$.

Якщо позначити вектор $\vec{\eta} = (\eta_1, \ldots, \eta_n)$, то за визначенням його густина розподілу дорівнює

$$p_{\vec{\eta}}(x_1, \ldots, x_n) = \frac{1}{\prod\limits_{j=1}^n \left[2\pi(t_j - t_{j-1})\right]^{1/2}} \prod_{j=1}^n \left[-\frac{x_j^2}{2(t_j - t_{j-1})}\right] .$$

Покладемо $t_0 = 0$, тоді в силу умови 3) маємо:

$$\xi(t_1) = \eta_1\,,\ \xi(t_2) = \xi(t_1) + \eta_2 = \eta_1 + \eta_2\,, \ldots,\ \xi(t_n) = \eta_1 + \ldots + \eta_n\,,$$

тобто випадкові вектори $\vec{\eta} = (\eta_1, \ldots, \eta_n)$ та $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ зв'язані невиродженим перетворенням із визначником, що дорівнює одиниці. Тому $\vec{\xi}$ також є нормальним та має густину $(x_0 = 0,\ t_0 = 0)$

$$p_{\vec{\xi}}(x_1, \ldots, x_n) = \frac{1}{\prod\limits_{j=1}^n \left[2\pi(t_j - t_{j-1})\right]^{1/2}} \prod_{j=1}^n \left[-\frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})}\right] .$$

Розглянемо дискретну одновимірну модель *броунівського руху* частинки, коли частинка змінює свої положення стрибками через певні рівні моменти часу Δt. Нехай стрибок дорівнює Δx. Тоді, якщо в момент часу t частинка мала координату x, то в момент часу $t+\Delta t$ вона з рівними ймовірностями буде знаходитись в точках $x-\Delta x$ або $x+\Delta x$. Зміщення частинки в момент часу $t+s$ від початкового положення в момент часу $t=0$ можна представити у вигляді

$$\xi(t+s)=[\xi(t+s)-\xi(s)]+[\xi(s)-\xi(0)]\,,\quad \xi(0)=0\,, \tag{10.7}$$

причому члени в квадратних дужках правої частини (10.7) незалежні і мають однакові розподіли. Тоді

$$D\xi(t+s)=D\,[\xi(t+s)-\xi(s)]+D\,[\xi(s)-\xi(0)]=D\xi(t)+D\xi(s)\,. \tag{10.8}$$

Із виразу (10.8) випливає, що дисперсія є лінійною функцією t, тобто

$$D\xi(t)=\sigma^2 t,\quad 0\le t<\infty\,,$$

де σ^2 - *коефіцієнт дифузії.*

За час t частинка здійснить $n=\dfrac{t}{\Delta t}$ стрибків. Позначимо ξ_k випадкову величину, що з ймовірностями $\dfrac{1}{2}$ приймає значення $\pm\Delta x$. Тоді

$$\xi(t)=\sum_{k=1}^{t/\Delta t}\xi_k,\; D\xi(t)=\sum_{k=1}^{t/\Delta t}D\xi_k=\sum_{k=1}^{t/\Delta t}\left[\frac{1}{2}(\Delta x)^2+\frac{1}{2}(-\Delta x)^2\right]=\frac{(\Delta x)^2 t}{\Delta t}\,. \tag{10.9}$$

Порівнюючи (10.8) та (10.9), отримаємо

$$\sigma^2=\frac{(\Delta x)^2}{\Delta t}\,.$$

Введемо випадкову величину

$$\xi_n(t)=\frac{\xi(t)}{\sqrt{D\xi(t)}}=\frac{\xi(t)}{\sigma\sqrt{t}}=\frac{1}{\sigma\sqrt{t}}\sum_{k=1}^{t/\Delta t}\xi_k\,.$$

Оскільки $M\xi_n(t)=0$ та $D\xi_n(t)=1$, то внаслідок центральної граничної теореми при $\Delta t\to 0$ та $\Delta x\to 0$, так що $\dfrac{(\Delta x)^2}{\Delta t}=\sigma^2$ випадкова величина $\xi_n(t)$ прямує за розподілом до нормального закону:

$$P\left\{x_1\le\frac{\xi(t)}{\sqrt{D\xi(t)}}\le x_2\right\}=\frac{1}{\sqrt{2\pi}}\int\limits_{x_1}^{x_2}e^{-\frac{x^2}{2}}dx\,,$$

і в силу однорідності процесу броунівського руху

$$P\left\{x_1\le\frac{\xi(t+s)-\xi(s)}{\sqrt{D\xi(t)}}\le x_2\right\}=\frac{1}{\sqrt{2\pi}}\int\limits_{x_1}^{x_2}e^{-\frac{x^2}{2}}dx\,,$$

тобто *броунівський рух - це вінеровський процес.*

10.3. Марковські процеси

Означення. Випадковий процес $\xi(t)$ називається *марковським*, якщо для будь-якого t_1 при відомому значенні $\xi(t_1)$ випадкові величини $\xi(t)$ з аргументами $t > t_1$ не залежать від випадкових величин $\xi(s)$ з аргументами $s < t_1$ (*процеси без наслідків*).

Нехай час є неперервним, а кількість станів скінченною: $\Omega = (\omega_1, \ldots, \omega_N)$. Цей процес відрізняється від ланцюгів Маркова тим, що перехід системи від стану ω_i до стану ω_j може відбуватися в довільний момент часу. Визначимо *ймовірності переходу* $p_{ij}(s,t)$

$$p_{ij}(s,t) = P\left\{\xi(t) = \omega_j | \xi(s) = \omega_i\right\}, \quad p_{ij}(s,t) \geq 0\,.$$

Очевидно, що

$$\sum_{j=1}^{N} p_{ij}(s,t) = 1,\ s \leq t,\ p_{ii}(t,t) = 1,\ p_{ij}(s,t) = 1,\ i \neq j\,. \tag{10.10}$$

Позначимо через a_i початковий розподіл ймовірностей в момент часу $t = t_0$:

$$a_i \geq 0, \quad \sum_{i=1}^{N} a_i = 1\,,$$

та через $p_j(t)$ - *абсолютну ймовірність*, тобто ймовірність того, що у момент часу $t \geq t_0$ буде заселеним стан ω_j. Внаслідок формули повної ймовірності є справедливими такі співвідношення

$$p_j(t) = \sum_{k=1}^{N} a_k p_{kj}(t_0, t), \quad t \geq t_0 \tag{10.11}$$

і

$$p_{ij}(s,t) = \sum_{k=1}^{N} p_{ik}(s,t_1) p_{kj}(t_1,t), \quad s < t_1 < t\,. \tag{10.12}$$

Теорема 25. (Колмогорова) *Нехай перехідні ймовірності $p_{ij}(s,t)$ мають частинні похідні по t і s. Тоді при $s \leq t$ мають місце співвідношення*

$$\frac{\partial p_{ij}(s,t)}{\partial t} = \sum_{k=1}^{N} p_{ik}(s,t) A_{kj}(t)\,, \tag{10.13}$$

$$\frac{\partial p_{ij}(s,t)}{\partial s} = -\sum_{k=1}^{N} A_{ik}(s) p_{kj}(s,t)\,, \tag{10.14}$$

де

$$\begin{gathered} A_{ik}(t) = \left[\frac{\partial p_{ik}(t,u)}{\partial u}\right]\Bigg|_{u=t},\ A_{ik}(t) \geq 0,\ i \neq k,\ A_{kk}(t) \leq 0\,, \\ \sum_{k=1}^{N} A_{ik}(t) = 0\,,\ i = \overline{1,N}\,. \end{gathered} \tag{10.15}$$

◁ Внаслідок (10.12) при $t > s$ маємо

$$
\begin{aligned}
&\frac{\partial p_{ij}(s,t)}{\partial t} = \lim_{\Delta\to 0}\frac{p_{ij}(s,t+\Delta)-p_{ij}(s,t)}{\Delta} = \\
&= \lim_{\Delta\to 0}\frac{1}{\Delta}\left[\sum_{k=1}^{N} p_{ik}(s,t)p_{kj}(t,t+\Delta) - p_{ij}(s,t)\right] = \\
&= \lim_{\Delta\to 0}\left[\sum_{k=1,\,k\neq j}^{N} p_{ik}(s,t)\frac{p_{kj}(t,t+\Delta)}{\Delta} + p_{ij}(s,t)\frac{p_{jj}(t,t+\Delta)-1}{\Delta}\right].
\end{aligned}
\tag{10.16}
$$

Розглянемо вираз в останніх квадратних дужках

$$
\begin{aligned}
&\lim_{\Delta\to 0}\frac{p_{jj}(t,t+\Delta)-1}{\Delta} = \left.\frac{\partial p_{jj}(t,u)}{\partial t}\right|_{u=t} \equiv A_{jj}(t)\,, \\
&\lim_{\Delta\to 0}\frac{p_{kj}(t,t+\Delta)}{\Delta} = \left.\frac{\partial p_{kj}(t,u)}{\partial t}\right|_{u=t} \equiv A_{kj}(t)\,, \quad k\neq j\,.
\end{aligned}
\tag{10.17}
$$

Підставляючи цей вираз в (10.16), отримаємо першу систему рівнянь (10.13). Співставляючи (10.17) та (10.10), приходимо до висновку, що внаслідок нерівностей $0 \le p_{ij}(t,u) \le 1$ є обмеження: $A_{ik} \ge 0$ при $i \neq k$, $A_{kk} \le 0$, $\sum\limits_{k=1}^{N} A_{ik}(t) = 0$, і (10.15) також доведено. Далі, при $t > s$ внаслідок (10.12) маємо

$$
\begin{aligned}
&\frac{\partial p_{ij}(s,t)}{\partial s} = \lim_{\Delta\to 0}\frac{p_{ij}(s+\Delta,t)-p_{ij}(s,t)}{\Delta} = \\
&= \lim_{\Delta\to 0}\frac{1}{\Delta}\left[p_{ij}(s+\Delta,t) - \sum_{k=1}^{N} p_{ik}(s,s+\Delta)p_{kj}(s+\Delta,t)\right] = \\
&= -\lim_{\Delta\to 0}\left[\frac{p_{ii}(s,s+\Delta)-1}{\Delta}p_{ij}(s+\Delta,t) + \sum_{k=1,\,k\neq i}^{N}\frac{p_{ik}(s,s+\Delta)}{\Delta}p_{kj}(s+\Delta,t)\right].
\end{aligned}
$$

Підставляючи у ці рівняння (10.17) отримаємо другу систему рівнянь (10.14). ▷
Рівняння (10.13) називаються *прямою*, а рівняння (10.14) *зворотною системою рівнянь Колмогорова.*

Фізичний зміст $A_{ij}(t)$ полягає в тому, що $A_{ij}(t)dt$ є ймовірність переходу із стану ω_i в стан ω_j в інтервалі часу $t, t+dt$ (тобто $A_{ij}(t)$ - швидкість переходу). Якщо $A_{ij}(t)$ є неперервними, то $p_{ij}(s,t)$ представляють собою єдиний розв'язок (10.11) із початковими умовами

$$
p_{ii}(s,s) = 1, \quad p_{ij}(s,s) = 0,\, i \neq j\,.
$$

Якщо a_i - початковий розподіл ймовірностей і $p_i(t)$ - абсолютні ймовірності, то

$$
p_i(t) = \sum_{j=1}^{N} a_j p_{ji}(t_0,t)\,.
$$

Продиференцюємо цей вираз і використаємо (10.10)

$$\begin{aligned} \frac{dp_i(t)}{dt} &= \sum_{j=1}^{N} a_j \sum_{k=1}^{N} p_{jk}(t_0, t) A_{kj}(t) = \\ &= \sum_{k=1}^{N} A_{kj}(t) \sum_{j=1}^{N} a_j p_{jk}(t_0, t) = \sum_{k=1}^{N} A_{kj}(t) p_j(t) \end{aligned} \tag{10.18}$$

з початковими умовами $p_i(t_0) = a_i, \quad i = \overline{1, N}$.

Приклад 53. (Двостороння реакція.) *Система може знаходитися у двох станах* ω_1 *та* ω_2 *(*ω_1 *- частинка, що не розпалась,* ω_2 *- розпалась). Знайти ймовірності заселення цих станів в залежності від часу.*

$\triangleleft$ Нехай є можливим як процес розпаду - перехід із стану ω_1 в стан ω_2 з ймовірністю $\alpha\Delta t + o(\Delta t)$ за час Δt, так і процес відновлення - перехід із стану ω_2 в стан ω_1 з ймовірністю $\beta\Delta t + o(\Delta t)$ за час Δt. Таким чином, у цьому випадку $A_{12} = \alpha$, $A_{21} = \beta$, і значить, $A_{11} = -\alpha$, $A_{22} = -\beta$. Рівняння (10.18) набудуть вигляду

$$\frac{dp_1(t)}{dt} = -\alpha p_1(t) + \beta p_2(t)\,, \quad \frac{dp_2(t)}{dt} = \alpha p_1(t) - \beta p_2(t)\,. \tag{10.19}$$

Нехай заданий початковий розподіл $a_1 = 1$, $a_2 = 0$. Внаслідок умови $\sum_k p_k(t) = 1$ випливає зв'язок $p_2(t) = 1 - p_1(t)$, який ми підставимо у перше з рівнянь (10.19),

$$\frac{dp_1(t)}{dt} = -(\alpha + \beta) p_1(t) + \beta\,.$$

Звідки з урахуванням початкового розподілу знайдемо

$$\begin{aligned} p_1(t) &= e^{-(\alpha+\beta)t} + \frac{\beta}{\alpha+\beta}(1 - e^{-(\alpha+\beta)t})\,, \\ p_2(t) &= \frac{\alpha}{\alpha+\beta}(1 - e^{-(\alpha+\beta)t})\,. \end{aligned}$$

При $t \to \infty$ маємо $p_1(t) \to \dfrac{\beta}{\alpha+\beta}$, $p_1(t) \to \dfrac{\alpha}{\alpha+\beta}$, тобто, процес є ергодичним. Якщо відновлення неможливо (наприклад, радіоактивний розпад), то $\beta = 0$ і

$$p_1(t) = e^{-\alpha t}\,, \quad p_2(t) = 1 - e^{-\alpha t}\,.$$

$\triangleright$

Розглянемо тепер марковські процеси $\xi(t)$ із неперервною множиною станів. Нехай для кожного набору моментів часу $t_1, \ldots, t_n \in T$ n-вимірна випадкова величина $\xi(t_1), \ldots, \xi(t_n)$ має n-вимірну густину ймовірності $p_n(t_1, x_1; \ldots; t_n, x_n)$. Ця густина має очевидні властивості: 1) $p_n(t_1, x_1; \ldots; t_n, x_n)$ є симетричною відносно будь-яких перестановок пар аргументів (t_i, x_i), оскільки $p_n(t_1, x_1; \ldots; t_n, x_n)\, dx_1 \ldots dx_n$ виражає ймовірність сумісного здійснення подій $x_i \leq \xi(t_i) \leq x_i + dx_i$, $i = \overline{1, n}$, і, таким чином, не залежить від порядку їх переліку;

2) усі скінченовимірні густини p_n для різних n повинні бути узгодженими у тому розумінні, що густина довільного k-вимірного розподілу при $k < n$ визначається за допомогою n-вимірного розподілу:

$$p_k(t_1, x_1; \ldots; t_k, x_k) = \tag{10.20}$$

$$= \underbrace{\int_{-\infty}^{\infty} \ldots \int}_{n-k} p_n(t_1, x_1; \ldots; t_k, x_k; t_{k+1}, x_{k+1}; \ldots; t_n, x_n) dx_{k+1} \ldots dx_n \,.$$

Відповідно до визначення густини ймовірності умовного розподілу (див. розділ (4.5.)) справедлива рівність

$$p_n(t_1, x_1; \ldots; t_n, x_n) = \tag{10.21}$$

$$= p_{n-1}(t_1, x_1; \ldots; t_{n-1}, x_{n-1}) q_n(t_n, x_n | t_1, x_1; \ldots; t_{n-1}, x_{n-1}) \,.$$

Ми говоримо, що процес *марковський*, коли його ймовірнісні властивості в момент часу t_n визначаються станом системи в момент часу t_{n-1} і не залежать від того, як протікав процес в попередні моменти часу, внаслідок чого

$$q_n(t_n, x_n | t_1, x_1; \ldots; t_{n-1}, x_{n-1}) = q(t_n, x_n | t_{n-1}, x_{n-1}) \,. \tag{10.22}$$

Умовна густина ймовірності $q(t, x|\tau, y)$ називається *перехідною густиною ймовірності.*

Співставляючи (10.21) та (10.22), можна знайти

$$\begin{aligned} p_n(t_1, x_1; ..; t_n, x_n) &= p_{n-1}(t_1, x_1; ..; t_{n-1}, x_{n-1}) q(t_n, x_n | t_{n-1}, x_{n-1}) = \\ &= p_1(t_1, x_1) q(t_2, x_2 | t_1, x_1) \ldots q(t_n, x_n | t_{n-1}, x_{n-1}) \,. \end{aligned} \tag{10.23}$$

Рівність (10.23) означає, що для того, щоб задати n-вимірну густину ймовірності марковського процесу достатньо знати лише дві функції: одновимірну густину $p_1(t, x)$ і перехідну густину ймовірностей $q(t, x|\tau, y)$.

Для довільних моментів часу $t_0 < \tau < t$, $t_0, \tau, t \in T$ згідно з (10.23) мають місце співвідношення

$$p_3(t_0, x_0; \tau, y; t, x) = p_1(t_0, x_0) q(\tau, y | t_0, x_0) q(t, x | \tau, y) \,, \tag{10.24}$$

та

$$p_2(t_0, x_0; t, x) = p_1(t_0, x_0) q(t, x | t_0, x_0) \,. \tag{10.25}$$

Тоді внаслідок умови узгодження (10.20)

$$p_2(t_0, x_0; t, x) = \int_{-\infty}^{\infty} p_3(t_0, x_0; \tau, y; t, x) dy \,. \tag{10.26}$$

Підставляючи (10.24) та (10.25) в (10.26) отримаємо рівняння

$$q(t,x|t_0,x_0) = \int\limits_{-\infty}^{\infty} q(t,x|\tau,y)q(\tau,y|t_0,x_0)dy\,. \tag{10.27}$$

Це рівняння називається *рівнянням Смолуховського* (або *Колмогорова-Чепмена*) і є основним в теорії неперервних марковських процесів.

Для *однорідного* марковського процесу густина ймовірності переходу залежить тільки від різниці моментів часу $q(t,x|\tau,y) = q(x|t-\tau,y)$. У цьому випадку рівняння Смолуховського (10.27) має вигляд

$$q(x|t-t_0,x_0) = \int\limits_{-\infty}^{\infty} q(x|t-\tau,y)q(y|\tau-t_0,x_0)dy\,.$$

Розглянемо більш детально *дифузійні процеси*, тобто поведінку частинки під дією випадкових поштовхів. Ці процеси характеризуються виконанням таких умов:
1)

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\int\limits_{-\infty}^{\infty}(x-y)q(t+\Delta,x|t,y)dx = A(y,t)\,. \tag{10.28}$$

Інтеграл у цьому виразі означає умовне середнє значення зміщення частинки за час Δ із фіксованої точки y, таким чином $A(y,t)$ визначає середню швидкість зміни стану в момент часу t. Величина $A(y,t)$ називається *коефіцієнтом зносу*.
2)

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\int\limits_{-\infty}^{\infty}(x-y)^2q(t+\Delta,x|t,y)dx = B(y,t)\,. \tag{10.29}$$

Тут інтеграл дає міру розкиду кінцевих точок x відносно початкової точки y. Умова 2) означає, що цей розкид росте при малих Δ пропорційно Δ, тобто по дифузійному закону. Коефіцієнт $\dfrac{1}{2}B(y,t)$ називається *коефіцієнтом дифузії*.
3)

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\int\limits_{-\infty}^{\infty}|x-y|^3q(t+\Delta,x|t,y)dx = 0\,.$$

Ця умова означає, що у малих проміжках часу ймовірність великих значень $|x-y|$ є малою. З фізичної точки зору саме ця умова дозволяє розглядати ξ як величину, що неперервно змінюється у часі, тобто як середнє значення на проміжках часу, що є набагато більшими за проміжок між двома послідовними випадковими поштовхами.

Якщо виконуються умови 1)-3) і функції $q(t,x|t_0,y_0)$, $A(x,t)$ та $B(x,t)$ задовольняють деяким додатковим умовам (які відносяться до їх диференційованості по своїх аргументах), то перехідна густина $q(t,x|t_0,y_0)$, як функція x і t задовольняє рівняння

$$\begin{aligned}\frac{\partial q(t,x|t_0,x_0)}{\partial t} &= -\frac{\partial}{\partial x}\left(A(x,t)q(t,x|t_0,x_0)\right)+ \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(B(x,t)q(t,x|t_0,x_0)\right)\,,\end{aligned} \tag{10.30}$$

і як функція x_0 і t_0 задовольняє рівняння

$$\begin{aligned}\frac{\partial q(t,x|t_0,x_0)}{\partial t_0} &= -A(x_0,t_0)\frac{\partial q(t,x|t_0,x_0)}{\partial x_0}- \\ &- \frac{1}{2}B(x_0,t_0)\frac{\partial^2 q(t,x|t_0,x_0)}{\partial x_0^2}\,.\end{aligned} \tag{10.31}$$

Рівняння (10.30) називається *рівнянням Ейнштейна-Фоккера-Планка* або *першим (прямим) рівнянням Колмогорова*. Це рівняння параболічного типу і його розв'язок повинен задовольняти умовам

$$q(t,x|t_0,x_0)\geq 0,\ \int\limits_{-\infty}^{\infty} q(t,x|t_0,x_0)dx=1,\ \int\limits_{-\infty}^{\infty} q(t,x|t_0,x_0)dx_0=1, \tag{10.32}$$

а також початковій умові

$$q(t,x|t_0,x_0)|_{t=t_0} = \delta(x-x_0)\,. \tag{10.33}$$

Умова (10.33) є наслідком рівняння (10.20).

Рівняння (10.31) називається *другим (оберненим) рівнянням Колмогорова* і його розв'язок повинен задовольняти тим же умовам (10.32) та (10.33). Це рівняння необхідно розв'язувати в обернену сторону за часом, для $t_0 \leq t$.

Ми не будемо тут доводити теореми, наслідками яких є рівняння (10.30) та (10.31), а замість цього розглянемо деякі приклади. Нехай у початковий момент часу t_0 є заданою густина розподілу ймовірності $p(t_0,x)$ випадкової величини $\xi(t_0)$. Тоді двовимірна густина ймовірності для довільного моменту часу $t \geq t_0$ і початкового моменту t_0 дорівнює

$$p_2(t,x|t_0,x_0) = p(t_0,x_0)q(t,x|t_0,x_0)\,,$$

а одновимірна густина процесу $\xi(t)$ в момент часу t згідно з (10.20) дорівнює

$$p_1(t,x) = \int\limits_{-\infty}^{\infty} p(t_0,x_0)q(t,x|t_0,x_0)dx_0\,. \tag{10.34}$$

Помножимо рівняння (10.30) на $p(t_0,x_0)$ і проінтегруємо по x_0. Внаслідок (10.34) густина $p_1(t,x)$ задовольняє таке ж рівняння

$$\frac{\partial p_1(t,x)}{\partial t} = -\frac{\partial}{\partial x}\left(A(x,t)p_1(t,x)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(B(x,t)p_1(t,x)\right)\,. \tag{10.35}$$

Розв'язок рівняння (10.35) повинен задовольняти умовам

$$p_1(t,x) \geq 0\,, \qquad \int\limits_{-\infty}^{\infty} p_1(t,x)dx = 1\,, \qquad p_1(t,x)|_{t=t_0} = p(t_0,x)\,. \tag{10.36}$$

Приклад 54. *Нехай марковський процес є однорідним у часі, тобто $q(t,x|t_0,x_0) = q(x|t-t_0,x_0)$. У цьому випадку згідно до (10.28) та (10.29) функції A і B не залежать від t. Нехай крім того одновимірна густина p_1 також не залежить від часу* **(стаціонарний марковський процес)**.

$\triangleleft$ Тоді рівняння (10.35) набуває вигляду

$$\frac{\partial}{\partial x}\left[A(x,t)p_1(t,x) - \frac{1}{2}\frac{\partial}{\partial x}\left(B(x,t)p_1(t,x)\right)\right]\,. \tag{10.37}$$

Якщо на межах зміни x (тобто області значень процесу $\xi(t)$) потік $Ap_1 - \dfrac{1}{2}\dfrac{\partial}{\partial x}(Bp_1)$ дорівнює нулеві, то внаслідок (10.37) він дорівнює нулеві всюди

$$A(x,t)p_1(t,x) - \frac{1}{2}\frac{\partial}{\partial x}\left(B(x,t)p_1(t,x)\right) = 0\,. \tag{10.38}$$

Інтегруючи рівняння (10.38), отримаємо

$$p_1(x) = \frac{c}{B(x)}\exp\left[\int\limits_0^x \frac{2A(y)}{B(y)}dy\right]\,, \tag{10.39}$$

де c - стала, що визначається із умови нормування (10.36). Фізичним прикладом реалізації такого стаціонарного розподілу є броунівський рух частинки над межею, що відбиває, і при наявності сили тяжіння. У цьому випадку $A = -mg$. Оскільки на межі виконується умова перетворення потоку в нуль, то є справедливим рівняння (10.38). Вираз (10.39) з постійними A і B дає

$$p_1(x) = c\,\exp\left[-\frac{2mgx}{B}\right]\,,$$

тобто ми отримали барометричну формулу. Відомо, що $B = 2kT$, де k - стала Больцмана, T - абсолютна температура. В результаті

$$p_1(x) = c\,\exp\left[-\frac{mgx}{kT}\right]\,.$$

$\triangleright$

Приклад 55. *Нехай марковський процес є однорідним по координаті: $q(t,x|t_0,x_0) = q(t,x-x_0|t_0)$.*

$\triangleleft$ У цьому випадку згідно з (10.28) та (10.29) функції A і B залежать лише від t і рівняння (10.30) набуває вигляду

$$\frac{\partial q(t, x-x_0|t_0)}{\partial t} = -A(t)\frac{\partial q}{\partial x} + \frac{B(t)}{2}\frac{\partial^2 q}{\partial x^2}\,.$$

Зробимо заміну змінних $(x,t) \to (y,\tau)$, де

$$y = x - x_0 - \int\limits_{t_0}^{t} A(s)ds\,, \qquad \tau = \int\limits_{t_0}^{t} B(s)ds\,,$$

і перейдемо до рівняння теплопровідності

$$\frac{\partial q}{\partial \tau} = \frac{1}{2}\frac{\partial^2 q}{\partial y^2}\,.$$

Його розв'язок, що відповідає умовам (10.32) та (10.33), має вигляд

$$q(\tau, y) = \frac{1}{\sqrt{2\pi\tau}}\exp\left[-\frac{y^2}{2\tau}\right]\,,$$

або в старих змінних

$$q(t, x-x_0|t_0) = \left(2\pi\int\limits_{t_0}^{t} B(s)ds\right)^{-\frac{1}{2}} \exp\left[-\frac{(x - x_0 - \int_{t_0}^{t} A(s)ds)^2}{2\int_{t_0}^{t} B(s)ds}\right]\,.$$

Нехай, зокрема, коефіцієнт зносу $A(t) = 0$, а величина $B(t) = B$ є сталою. Тоді

$$q(t, x-x_0|t_0) = (2\pi B(t-t_0))^{-\frac{1}{2}}\exp\left[-\frac{(x-x_0)^2}{2B(t-t_0)}\right]\,. \tag{10.40}$$

Звідси випливає, що середнє зміщення частинки дорівнює нулеві, а середній квадрат зміщення частинки дорівнює $B(t-t_0)$, тобто зростає пропорційно часу. Цей результат уперше отримав Ейнштейн.

Хоча середнє значення змінної $\xi(t)$ дорівнює нулеві, середньоквадратичне значення необмежено зростає при $t \to \infty$. Це означає, що конкретні реалізації траєкторій $\xi(t)$ надзвичайно мінливі. Водночас, броунівський рух має ще деякі особливості. Так, оскільки він описується рівнянням дифузійного типу, то траєкторії $\xi(t)$ неперервні. В той же час вони є *недиференційовними*. Дійсно, відповідно до (10.40), можна обчислити ймовірність події $\left\{\left|\frac{\xi(t+h)-\xi(t)}{h}\right| > k\right\}$:

$$P\left\{\left|\frac{\xi(t+h)-\xi(t)}{h}\right| > k\right\} = 2\int\limits_{kh}^{\infty}(2\pi Bh))^{-\frac{1}{2}}\exp\left[-\frac{x^2}{2Bh}\right]\,,$$

і ця ймовірність перетворюється в одиницю коли $h \to 0$. Це означає, що є неважливим яке значення k обрати, – все одно величина $\left|\frac{\xi(t+h)-\xi(t)}{h}\right|$ майже напевно буде більша цього значення, тобто похідна по часу у будь-якій точці майже напевно нескінченна. Звичайно, що швидкість броунівської частинки фізично не може бути нескінченною, і тому треба розглядати більш реалістичні моделі броунівського руху (наприклад, *процес Орнштейна-Уленбека*). $\triangleright$

Зміст