МІНІСТЕРСТВО ОСВІТИ УКРАЇНИ
Івано-Франківський державний технічний
університет нафти і газу

ТЕОРІЯ ЙМОВІРНОСТЕЙ, ВИПАДКОВІ ПРОЦЕСИ ТА МАТЕМАТИЧНА СТАТИСТИКА

КОНСПЕКТ ЛЕКЦІЙ

ЧАСТИНА 1

ТЕОРІЯ ЙМОВІРНОСТЕЙ

МВ 02070855 – 475 – 99

Затверджено на засіданні
кафедри вищої математики
Протокол № 8 від 19.05.99

Всі цитати, цифровий і фактичний матеріал перевірені.
Зауваження рецензента враховані.
М.М.Осипчук

м. Івано-Франківськ
1999 р.

Конспект лекцій призначений для студентів всіх спеціальностей денної та заочної форм навчання, які вивчають дисципліни "Теорія ймовірностей та математична статистика"і "Теорія ймовірностей та випадкові процеси", а також "Вища математика"(розділ "Теорія ймовірностей").

Автор: канд. фіз.-мат. наук, доцент Осипчук М.М.

Відповідальний за випуск завідувач кафедри вищої математики, доцент Горгула В.І.

Погоджено з навчально-методичними об'єднаннями спеціальностей:
7.091401 Л.М. Заміховський

7.092501 Г.Н. Семенцов

Нормоконтролер О.Г. Гургула
Коректор Н.Ф. Будуйкевич
Член експертно-рецензійної
комісії університету А.І.Волобуєв

Вступ

Конспект лекцій призначений для самостійного вивчення студентами всіх спеціальностей денної та заочної форм навчання дисциплін: "Теорія ймовірностей та математична статистика"і "Теорія ймовірностей та випадкові процеси".

Перша частина конспекту лекцій присвячена теорії ймовірностей і включає в себе такі базові для вивчення всього курсу розділи:

- випадкові події;
- випадкові величини;
- граничні теореми теорії ймовірностей.

Одна лекція, присвячена вивченню елементів комбінаторики, виділена окремо в додатку через її деяку відстороненість від загальної схеми згаданих дисциплін. Питання, розглянуті в цій лекції, можуть бути відомі студентам з курсу середньої школи чи з інших математичних дисциплін, що вивчаються у вузі.

Після кожного розділу дано вправи, які можна використати для контролю засвоєння відповідного матеріалу.

Нумерація формул, таблиць, рисунків в кожному розділі окрема, а нумерація означень, теорем, лем, наслідків окрема в кожній лекції.

Доведення кожного твердження починається словом "Доведення"і закінчується значком .

Розділ 1

Випадкові події та ймовірності

1.1 Стохастичний експеримент та ймовірнісний простір

1.1.1 Стохастичний експеримент, елементарні події, події

Вихідним поняттям теорії ймовірностей є поняття стохастичного експерименту, випадкової події та ймовірності випадкової події. *Стохастичним (випадковим) експериментом* будемо називати експеримент (деяку дію), результат якого неможливо точно передбачити наперед. При цьому будемо розглядати тільки такі експерименти, які можна повторювати при незмінних умовах будь-яку кількість раз. Будь-який результат стохастичного експерименту, який може бути зафіксованим (наблюдатися), будемо називати *випадковою подією* [1](подією). Серед випадкових подій виділимо ті з них, які не можуть відбуватися одночасно і не можуть бути розкладені на простіші випадкові події. Такі події будемо називати *елементарними подіями*.

Таким чином, з кожним стохастичним експериментом можна пов'язати деяку множину Ω елементарних подій (*простір елементарних подій*). Кожну подію, яка може настати в результаті цього експерименту, можна тому розглядати як деяку підмножину простору елементарних подій.

Елементарні події будемо позначати буквами ω або ω_i ($\omega \in \Omega$ або $\omega_i \in \Omega$), випадкові події будемо позначати великими буквами латинського алфавіту: A, B, і т. д. Виняток становлять дві події: Ω — вірогідна подія та $\emptyset$ — неможлива подія.

Будемо говорити, що подія A настала (відбулася), якщо настала яка-небуть елементарна подія $\omega \in A$, при цьому вважатимемо, що елементарна подія ω сприяє події A.

Приклад 1. Експеримент полягає в трикратному підкиданні монети і фіксації сторони монети, якою вона випаде кожного разу. В такому експерименті є всього 8 найпростіших результатів (елементарних подій). Тобто

$$\Omega = \{\text{ггг, ггц, гцг, цгг, гцц, цгц, ццг, ццц}\}.$$

Нехай подія A полягає в тому, що випаде принаймні один герб, а подія B — в тому, що другим випаде герб. Тому $A = \{\text{ггг, ггц, гцг, цгг, гцц, цгц, ццг}\}$, $B = \{\text{ггг, ггц, цгг, цгц}\}$.

Приклад 2. Експеримент полягає в киданні точки в деякий квадрат. Елементарною подією слід в цьому випадку вважати попадання в деяку фіксовану точку квадрата. Тому простір елементарних подій можна задати як множину точок квадрата.

Нехай подія A полягає в попаданні точки у вписаний в даний квадрат круг. Отже, A задається кругом, вписаним в даний квадрат.

---

[1] Пізніше дамо більш точне означення випадкової події.

Зауваження. Потрібно відмітити, що в першому прикладі простір елементарних подій складається із скінченної кількості елементів, а в другому — із нескінченної (незліченної).

1.1.2 Дії над подіями

Розглянемо тепер можливі дії над подіями та відношення, в яких ці події можуть бути. Дамо такі означення.

Означення. Будемо говорити, що подія A є частковим випадком події B (або B є наслідком A), якщо множина елементарних подій, що сприяють події A є підмножиною елементарних подій, що сприяють події B. Позначається це так: $A \subset B$ або $B \supset A$

Якщо $A \subset B$ і $B \subset A$, то кажуть, що події A і B співпадають, тобто $A = B$.

Означення. Сумою (або об'єднанням) двох подій A і B називається подія, яка полягає в тому, що відбудеться або подія A, або подія B. Позначається сума подій так: $A + B$ або $A \cup B$.

Тому сумі двох подій A і B сприяють ті і тільки ті елементарні події, які сприяють хоча б одній з подій A чи B.

Означення. Добутком (або перерізом) двох подій A і B називається подія, яка полягає в тому, що відбудеться і подія A, і подія B. Позначається добуток подій так: $A \cdot B$ або $A \cap B$.

Отже, добутку двох подій A і B сприяють ті і тільки ті елементарні події, які сприяють обидвом подіям A і B.

Означення. Протилежною подією до події A називається подія, яка полягає в тому, що подія A не відбудеться. Позначається протилежна подія так: $\overline{A}$.

Тобто, події $\overline{A}$ сприяють ті і тільки ті елементарні події, які не сприяють події A.

Крім розглянутих дій над подіями можна виділити ще й такі:

1. різниця подій $A \setminus B = A \cdot \overline{B}$; тому $\overline{A} = \Omega \setminus A$;

2. симетрична різниця подій $A \nabla B = (A \setminus B) + (B \setminus A)$.

Наведемо найважливіші властивості дій над подіями:

1. $A + B = B + A$, $A \cdot B = B \cdot A$ (комутативність додавання та множення);

2. $A + (B + C) = (A + B) + C$, $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (асоціативність додавання та множення);

3. $(A + B) \cdot C = A \cdot C + B \cdot C$ (дистрибутивність множення відносно додавання);

4. $A \cdot B + C = (A + C) \cdot (B + C)$ (дистрибутивність додавання відносно множення);

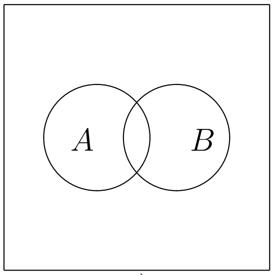
5. $\overline{A + B} = \overline{A} \cdot \overline{B}$, $\overline{A \cdot B} = \overline{A} + \overline{B}$ (закони де Моргана);

6. $A + \overline{A} = \Omega$, $A \cdot \overline{A} = \emptyset$, $A = \overline{\overline{A}}$.

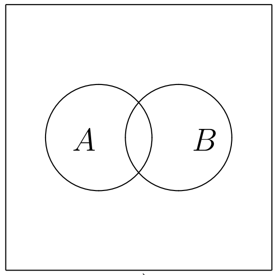
Легко довести ці та інші властивості дій над подіями з допомогою, наприклад, діаграм Ейлера-Венна. На цих діаграмах простір елементарних подій Ω зображається квадратом, а події — кругами.

Розглянемо властивість 5 (рис.1.1)(правильність інших властивостей встановлюється аналогічно):

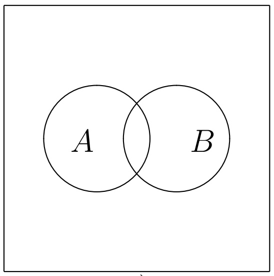
1. заштрихована подія $\overline{A + B}$;



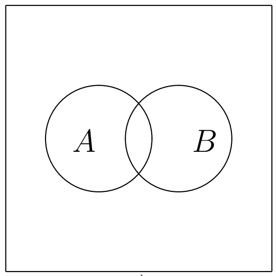
а)



б)



в)



г)

Рис. 1.1:

2. подвійною штриховкою заштрихована подія $\overline{A} \cdot \overline{B}$;
3. заштрихована подія $\overline{A \cdot B}$;
4. заштрихована (будь-якою штриховкою) подія $\overline{A} + \overline{B}$.

Означення. Події A і B називаються несумісними, якщо $A \cdot B = \emptyset$. В іншому випадку ці події сумісні.

Очевидно, що події A і $\overline{A}$ несумісні.

Нехай $\mathcal{A}$ — деяка множина подій (система підмножин Ω).

Означення. Множина подій $\mathcal{A}$ називається алгеброю, якщо:

1. $\Omega \in \mathcal{A}$;
2. з того, що $A,\ B \in \mathcal{A}$, випливає, що $A + B \in \mathcal{A}$;
3. з того, що $A \in \mathcal{A}$, випливає, що $\overline{A} \in \mathcal{A}$.

З цього означення легко слідує, що $\emptyset \in \mathcal{A}$, бо $\emptyset = \overline{\Omega}$, і якщо $A,\ B \in \mathcal{A}$, то $A \cdot B \in \mathcal{A}$, бо $A \cdot B = \overline{\overline{A} + \overline{B}}$.

Означення. Множина подій $\mathcal{F}$ називається σ-алгеброю (сіґма-алгеброю), якщо вона є алгеброю і для будь-якої послідовності подій $\{A_n,\ n \geq 1\}$ з $\mathcal{F}$ їх сума $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$.

Очевидно також, що [2] $\bigcap\limits_{n=1}^{\infty} A_n = \overline{\bigcup\limits_{n=1}^{\infty} \overline{A_n}} \in \mathcal{F}$, якщо $\{A_n,\ n \geq 1\} \subset \mathcal{F}$.

Надалі для кожного стохастичного експеримента будемо визначати разом з простором елементарних подій Ω також і σ-алгебру подій $\mathcal{F}$, і тільки елементи σ-алгебри $\mathcal{F}$ будемо називати *подіями*.

[2]Значками $\bigcup\limits_{n=1}^{\infty}$ та $\bigcap\limits_{n=1}^{\infty}$ позначені, відповідно, сума і добуток подій, записаних після них.

1.1.3 Ймовірність

В теорії ймовірностей та її застосуваннях майже завжди виникає потреба оцінювати події (що відносяться до одного чи різних стохастичних експериментів) за ступенем їх "достовірності"як, наприклад, подія Ω відбувається завжди (вірогідна подія), а подія $\emptyset$ не відбувається ніколи (неможлива подія). Для цього вводиться поняття ймовірності події.

Означення. Ймовірністю (ймовірнісною мірою) називається числова функція $\mathbf{P}(A)$, яка визначена на σ-алгебрі подій $\mathcal{F}$ та має такі властивості:

1. $\mathbf{P}(A) \geq 0$ для всіх $A \in \mathcal{F}$;
2. $\mathbf{P}(\Omega) = 1$;
3. якщо події A_1, A_2, ... попарно несумісні (тобто $A_i \cdot A_j = \emptyset$ при всіх $i \neq j$), то

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

Означення. Трійку елементів $(\Omega, \mathcal{F}, \mathbf{P})$, де Ω — простір елементарних подій, $\mathcal{F}$ — σ-алгебра подій (підмножин Ω), $\mathbf{P}$ — ймовірнісна міра на $\mathcal{F}$, будемо називати ймовірнісним простором.

Зауваження. Властивості з означень σ-алгебри $\mathcal{F}$, та ймовірнісної міри $\mathbf{P}$ ще називають аксіомами теорії ймовірностей. Аксіоматичний підхід до побудови теорії ймовірностей запропонував в 1929 р. А.М.Колмогоров.

1.2 Властивості та приклади ймовірностей

1.2.1 Властивості ймовірності

Розглянемо деякі найпростіші властивості ймовірності, що випливають з означення.

Теорема 1. Для будь-якої події A $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.

Доведення. Розглянемо події A і $\overline{A}$. Оскільки $A + \overline{A} = \Omega$ і $A \cdot \overline{A} = \emptyset$, то $1 = \mathbf{P}(\Omega) = \mathbf{P}(A + \overline{A}) = \mathbf{P}(A) + \mathbf{P}(\overline{A})$. Звідси $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.

Наслідок. $\mathbf{P}(\emptyset) = 0$.

Доведення. Оскільки $\emptyset = \overline{\Omega}$, то за теоремою 1

$$\mathbf{P}(\emptyset) = 1 - \mathbf{P}(\Omega) = 1 - 1 = 0.$$

Теорема 2. (*монотонність ймовірності*) Якщо $A \subset B$, то $\mathbf{P}(A) \leq \mathbf{P}(B)$

Доведення. Подамо подію B у вигляді $B = A + \overline{A} \cdot B$. Оскільки $A \cdot (\overline{A} \cdot B) = \emptyset$, то $\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(\overline{A} \cdot B) \geq \mathbf{P}(A)$, бо $\mathbf{P}(\overline{A} \cdot B) \geq 0$.

Теорема 3. (*додавання ймовірності*) Якщо A і B довільні події, то $\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cdot B)$.

Доведення. Оскільки $A + B = A + \overline{A} \cdot B$ і $B = \overline{A} \cdot B + A \cdot B$ та $A \cdot (\overline{A} \cdot B) = \emptyset$ і $(\overline{A} \cdot B) \cdot (A \cdot B) = \emptyset$, то

$$\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(\overline{A} \cdot B) \text{ і}$$

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

Звідси $\mathbf{P}(\overline{A} \cdot B) = \mathbf{P}(B) - \mathbf{P}(A \cdot B)$ і

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

Теорема 4. Якщо $A \subset B$, то $\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A)$.

Доведення. Подамо подію B у вигляді $B = B \setminus A + A$. Очевидно, що $(B \setminus A) \cdot A = \emptyset$ і тому

$$\mathbf{P}(B) = \mathbf{P}(B \setminus A) + \mathbf{P}(A),$$

звідси і випливає твердження теореми.

Теорема 5. (*неперервність ймовірності*) Нехай дано послідовність подій A_1, $A_2, \ldots$ таку, що $A_1 \supset A_2 \supset \ldots$ і $A = \bigcap\limits_{n=1}^{\infty} A_n$. Тоді $\mathbf{P}(A) = \lim\limits_{n\to\infty} \mathbf{P}(A_n)$.

Доведення. Розкладемо кожну подію A_n на суму попарно несумісних подій $A_n = \bigcup\limits_{k=n}^{\infty} A_k \cdot \overline{A_{k+1}} \cup A$. Тому $\mathbf{P}(A_n) = \sum\limits_{k=n}^{\infty} \mathbf{P}(A_k \cdot \overline{A_{k+1}} + \mathbf{P}(A)$, звідки

$$\mathbf{P}(A_n) - \mathbf{P}(A) = \sum_{k=n}^{\infty} \mathbf{P}(A_k \cdot \overline{A_{k+1}}). \tag{1.1}$$

При $n = 1$ з цієї рівності випливає, що ряд
$\sum\limits_{k=1}^{\infty} \mathbf{P}(A_k \cdot \overline{A_{k+1}})$ збігається. Оскільки за формулою (1.1) $\mathbf{P}(A_n) - \mathbf{P}(A)$ є залишком цього ряду, то

$$\lim_{n\to\infty} \mathbf{P}(A_n) - \mathbf{P}(A) = 0.$$

Перейдемо тепер до розгляду деяких конкретних прикладів ймовірнісних просторів.

1.2.2 Класичне означення ймовірності

Припустимо, що деякому стохастичному експериментові відповідає ймовірнісний простір $(\Omega,\ \mathcal{F},\ \mathbf{P})$. Нехай простір елементарних подій Ω є скінченною множиною і $n = |\Omega|$. Нехай $\mathcal{F}$ — σ-алгебра всіх підмножин Ω. Крім того, нехай ймовірнісна міра $\mathbf{P}$ така, що всі елементарні події мають однакові ймовірності (рівноможливі). Тому $\mathbf{P}(\omega_i) = \frac{1}{n}$.

Нехай далі події A сприяють $m = |A|$ елементарних подій тобто $A = \bigcup\limits_{k=1}^{m} \omega_{i_k}$.

Оскільки $\mathbf{P}(\bigcup\limits_{k=1}^{m} \omega_{i_k}) = \sum\limits_{k=1}^{m} \mathbf{P}(\omega_{i_k}) = \frac{m}{n}$, то

$$\mathbf{P}(A) = \frac{m}{n}. \tag{1.2}$$

Рівність (1.2) і є класичним означенням ймовірності.

Приклад 1. Підкидається гральний кубик. Знайти ймовірність того, що число очок, яке випаде, кратне трьом.

Найпростіші результати описаного експерименту (елементарні події) $\omega_i = \{$випало число $i\}$, $i = \overline{1, 6}$.

Очевидно (якщо кубик правильний — виготовленний з однорідного матеріалу), що ймовірнісну міру потрібно визначити так, щоб $\mathbf{P}(\omega_i) = \frac{1}{6}$.

Описаній подій (позначимо її A) сприяють дві елементарні події $A = \{\omega_3, \omega_6\}$.

Тому $|A| = 2$ і $\mathbf{P}(A) = \frac{2}{6} = \frac{1}{3}$.

Зауважимо, що функція задана на $\mathcal{F}$ рівністю (1.2) задовольняє всі вимоги, які повинні виконуватись для ймовірнісної міри. Дійсно

1. $\mathbf{P}(A) = \frac{m}{n} \geq 0$;

2. $\mathbf{P}(\Omega) = \frac{n}{n} = 1$;

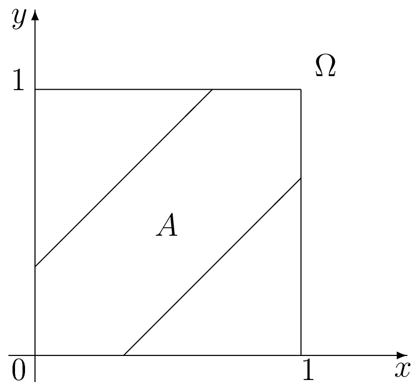


Рис. 1.2:

3. якщо $A \cdot B = \emptyset$ ($|A| = m$ і $|B| = k$), то

$$\mathbf{P}(A + B) = \frac{m+k}{n} = \frac{m}{n} + \frac{k}{n} = \mathbf{P}(A) + \mathbf{P}(B).$$

Зліченної адитивності від функції $\mathbf{P}$ вимагати не потрібно, бо $\mathcal{F}$ складається із скінченної кількості (2^n) підмножин.

1.2.3 Геометричне означення ймовірності

Розглянемо ситуацію коли стохастичному експерименту відповідає ймовірнісний простір $(\Omega,\ \mathcal{F},\ \mathbf{P})$ в якому простір елементарних подій Ω є нескінченним, навіть незліченним.

Нехай Ω є деякою множиною евклідового простору R^n ($n = 1, 2, 3$), σ-алгебра подій $\mathcal{F}$ — найменша σ-алгебра, що містить всі відкриті підмножини Ω. Як і в попередньому випадку, будемо вважати, що всі елементарні події рівноможливі. Проте ніякої ненульової ймовірності елементарним подіям приписати неможна, бо їх нескінченне число. Відштовхуючись від рівноможливості елементарних подій природно визначити ймовірність кожної події пропорційну її n-вимирній мірі (довжині при $n = 1$, площі при $n = 2$, об'єму при $n = 3$) $mes(A)$, як множини з R^n.

Тобто $\mathbf{P}(A) = C \cdot mes(A)$, а оскільки $1 = \mathbf{P}(\Omega) = = C \cdot mes(\Omega)$, то $C = \frac{1}{mes(\Omega)}$. Тому

$$\mathbf{P}(A) = \frac{mes(A)}{mes(\Omega)} \tag{1.3}$$

Рівність (1.3) і є геометричним означенням ймовірності.

Очевидно, що всі характеристичні властивості ймовірнісної міри виконуються, бо властивості 1) і 3) має сама міра mes, а $\mathbf{P}(\Omega) = \frac{mes(\Omega)}{mes(\Omega)} = 1$.

Приклад 2. (*задача про зустріч*) Дві особи А і Б домовились зустрітися в певному місці, причому кожна з них приходить туди незалежно від іншої у випадковий момент часу між 12 і 13 годиною.

Той, хто приходить першим, чекає 20 хвилин і йде. Знайти ймовірність того, що зустріч відбудеться.

Позначимо момент приходу А і Б відповідно через x і y. Тоді за простір елементарних подій природно взяти квадрат Ω у площині xOy

$$\Omega = \{(x; y) \in R^2 : 0 \le x \le 1,\ 0 \le y \le 1\}.$$

Згідно з умовою А і Б зустрінуться тоді і тільки тоді, коли $|x - y| \le \frac{1}{3}$ (різниця між моментами приходу не перевищує $\frac{1}{3}$ год.=20 хв.). Це означає, що зустрічі (подія A) відповідають точки квадрата, для яких виконується зазначена нерівність, тобто

$$A = \{(x; y) \in R^2 : 0 \le x \le 1,\ 0 \le y \le 1,\ |x - y| \le \frac{1}{3}\}$$

Побудуємо на координатній площині xOy області Ω і A (рис. 1.2).

Отже, шукана ймовірність, згідно з формулою (1.3)

$$\mathbf{P}(A) = \frac{mes(A)}{mea(\Omega)} = \frac{1 - 2 \cdot \frac{1}{2} \cdot \left(\frac{2}{3}\right)^2}{1} = \frac{5}{9}.$$

1.3 Умовні ймовірності. Незалежність випадкових подій

1.3.1 Умовна ймовірність

Нехай, деякому стохастичному експерименту відповідає ймовірнісний простір $(\Omega,\ \mathcal{F},\ \mathbf{P})$. Розглянемо дві події A і B, причому $\mathbf{P}(B) \neq 0$. Задамось питанням, чи впливає на ймовірність ("достовірність") події A інформація про те, що подія B відбулась, та як знайти цю ймовірність. Очевидно, що відповідь на першу частину питання повинна бути, взагалі кажучі, позитивною. Абсолютно зрозумілим є те, що якщо ми знаємо, що подія B відбулася, то простір елементарних подій необхідно звузити до множини, яка описує подію B. Ясно також, що замість σ-алгебри подій $\mathcal{F}$ слід взяти σ-алгебру $\mathcal{F}_B$, що є слідом $\mathcal{F}$ на події B. Тобто $\mathcal{F}_B = \{C \cap B; C \in \mathcal{F}\}$.

Легко довести, що множина подій $\mathcal{F}_B$ утворює σ-алгебру підмножин B.

Використовуючи ідеї побудови ймовірнісних мір, які ми розглядали вище, та поняття ймовірності як деякої міри ("ваги") події, дамо таке означення умовної ймовірності:

Означення. *Умовною ймовірністю* події A при умові, що відбулася подія B ($\mathbf{P}(B) \neq 0$), називається число $\mathbf{P}(A/B) = \frac{\mathbf{P}(A \cdot B)}{\mathbf{P}(B)}$.

Умовна ймовірність $\mathbf{P}(\cdot/B)$ має такі прості властивості:

1. $\mathbf{P}(A/B) \ge 0$ для всіх $A \in \mathcal{F}$;

2. $\mathbf{P}(\Omega/B) = \mathbf{P}(B/B) = 1$;

3. якщо події $A_1,\ A_2, \ldots$ попарно несумісні, то

$$\mathbf{P}\left(\bigcup_{n=1}^{\infty} A_n/B\right) = \frac{\mathbf{P}\left(\bigcup_{n=1}^{\infty} A_n \cdot B\right)}{\mathbf{P}(B)} = \frac{\sum_{n=1}^{\infty} \mathbf{P}(A_n \cdot B)}{\mathbf{P}(B)} = \\ = \sum_{n=1}^{\infty} \mathbf{P}(A_n/B),$$

бо події $\{A_n \cdot B : n \ge 1\}$ попарно несумісні.

Отже, $\mathbf{P}(\cdot/B)$ є ймовірнісною мірою як на $\mathcal{F}$ так і на $\mathcal{F}_B$. Оскільки для всіх $A \in \mathcal{F}$ таких, що $A \cdot B = \emptyset$, $\mathbf{P}(A/B) = 0$, то якщо відомо, що подія B відбулася, можна обмежитись умовним стохастичним експериментом, якому відповідає ймовірнісний простір $(B,\ \mathcal{F}_B,\ \mathbf{P}(\cdot/B))$.

Приклад 1. Двічі витягують без повернення по одній кульці з урни, яка містить n білих і m чорних кульок. Розглянемо дві події: A={друга кулька чорна}, B={перша кулька біла}.

Якщо подія B відбулась то в урні залишилось $n-1$ білих і m чорних кульок. Тому ймовірність події A при умові, що відбулася подія B дорівнює $\mathbf{P}(A/B)=\frac{m}{n-1+m}$. З іншого боку оскільки $\mathbf{P}(A\cdot B)=\frac{n\cdot m}{A^2_{n+m}}=\frac{n\cdot m}{(n+m)\cdot(n+m-1)}$, $\mathbf{P}(B)=\frac{n}{n+m}$, то $\mathbf{P}(A/B)=\frac{\mathbf{P}(A\cdot B)}{\mathbf{P}(B)}=\frac{m}{n+m-1}$.

Одержали два однакові результати, що є підтвердженням сказаного вище.

1.3.2 Теорема множення ймовірностей

Розглянемо деякі прості, але важливі властивості умовної ймовірності. Безпосередньо з означення випливає така теорема множення.

Теорема 1. Якщо $\mathbf{P}(B)\neq 0$, то

$$\mathbf{P}(A\cdot B)=\mathbf{P}(A/B)\cdot\mathbf{P}(B).$$

Через простоту доведення цієї тереми залишимо його для читача.

Очевидно, що якщо $\mathbf{P}(A)\neq 0$, $\mathbf{P}(B)\neq 0$, то

$$\mathbf{P}(A\cdot B)=\mathbf{P}(A/B)\cdot\mathbf{P}(B)=\mathbf{P}(B/A)\cdot\mathbf{P}(A).$$

Теорему 1 можна легко розповсюдити на будь-яку скінченну кількість подій.

Теорема 2. Якщо події A_1, $A_2,\ldots$, A_n такі, що $\mathbf{P}(A_1\cdot A_2\ldots A_{n-1})\neq 0$, то

$$\mathbf{P}(A_1\cdot A_2\ldots A_n)= \\ =\mathbf{P}(A_1)\cdot\mathbf{P}(A_2/A_1)\ldots\mathbf{P}(A_n/A_1\cdot A_2\ldots A_{n-1}). \tag{1.4}$$

Доведення. Очевидно, що умовні ймовірності в формулі (1.4) визначені, бо $\mathbf{P}(A_1\cdot A_2\ldots A_k)\neq 0$ при всіх
$k=1,2,\ldots,n-1$.

Дійсно оскільки

$$A_1\supset A_1\cdot A_2\supset\ldots\supset A_1\cdot A_2\ldots A_{n-1}$$

і ймовірнісна міра має властивість монотонності, то

$$\mathbf{P}(A_1\cdot A_2\ldots A_k)\geq\mathbf{P}(A_1\cdot A_2\ldots A_{n-1})>0$$

Формула (1.4) при $n=2$ правильна згідно теореми 1.

Нехай (1.4) правильна при деякому $n\in N$, тобто

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

Розглянемо

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

Тому за принципом математичної індукції рівність (1.4) має місце при кожному натуральному n.

1.3.3 Незалежні випадкові події

Незалежними,
природно, назвати випадкові події, ймовірності кожної з яких не залежать від того, відбулась чи ні інша подія. Як ми побачимо дальше, для незалежних подій умовні та безумовні ймовірності співпадають. Тому, виходячи з теореми множення, дамо таке означення незалежності двох подій:

Означення. Випадкові події A і B називаються незалежними, якщо

$$\mathbf{P}(A \cdot B) = \mathbf{P}(A) \cdot \mathbf{P}(B). \tag{1.5}$$

Звідси легко одержимо сказане вище.

Теорема 3. Якщо випадкові події A і B незалежні і
$\mathbf{P}(A) \neq 0$ та $\mathbf{P}(B) \neq 0$, то

$$\mathbf{P}(A/B) = \mathbf{P}(A) \quad \text{і} \quad \mathbf{P}(B/A) = \mathbf{P}(B).$$

.

Доведення. З означення умовної ймовірності

$$\mathbf{P}(A/B) = \frac{\mathbf{P}(A \cdot B)}{\mathbf{P}(B)} \quad \text{і} \quad \mathbf{P}(B/A) = \frac{\mathbf{P}(A \cdot B)}{\mathbf{P}(A)}$$

Оскільки A і B незалежні, то з (1.5) маємо

$$\mathbf{P}(A/B) = \frac{\mathbf{P}(A) \cdot \mathbf{P}(B)}{\mathbf{P}(B)} = \mathbf{P}(A) \quad \text{і}$$
$$\mathbf{P}(B/A) = \frac{\mathbf{P}(A) \cdot \mathbf{P}(B)}{\mathbf{P}(A)} = \mathbf{P}(B).$$

Доведемо ще одну властивість незалежних подій.

Теорема 4. Якщо події A і B незалежні, то незалежними є і події $\overline{A}$ і B та $\overline{A}$ і $\overline{B}$.

Доведення. Досить довести, що $\overline{A}$ і B незалежні.

Розглянемо $\mathbf{P}(\overline{A} \cdot B) = \mathbf{P}(B) - \mathbf{P}(A \cdot B)$, бо
$B = A \cdot B + \overline{A} \cdot B$ і $(A \cdot B) \cdot (\overline{A} \cdot B) = \emptyset$. Тому

$$\mathbf{P}(\overline{A} \cdot B) = \mathbf{P}(B) - \mathbf{P}(A \cdot B) =$$
$$= (1 - \mathbf{P}(A)) \cdot \mathbf{P}(B) = \mathbf{P}(\overline{A}) \cdot \mathbf{P}(B),$$

тобто, $\overline{A}$ і B незалежні.

Перейдемо тепер від двох подій до кількох.

Означення. Випадкові події $A_1, A_2, \ldots, A_n$ називаються незалежними в сукупності, якщо для будь-якого набору індексів $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ $(k = \overline{2, n})$

$$\mathbf{P}(A_{i_1} \cdot A_{i_2} \ldots A_{i_k}) = \mathbf{P}(A_{i_1}) \cdot \mathbf{P}(A_{i_2}) \ldots \mathbf{P}(A_{i_k}) \tag{1.6}$$

Якщо рівність (1.6) має місце при $k = 2$, то події $A_1, A_2, \ldots, A_n$ називають попарно незалежними.

Легко встановити, що із попарної незалежності, взагалі кажучи, незалежність в сукупності не випливає. В цьому переконує нас такий приклад. **Приклад 2.** (Бернштейна).

Нехай три грані однорідного тетраедра пофарбовані кожна в один з кольорів: синій, червоний, зелений. Четверта грань розфарбована так, що має всі три згадані кольори. Тетраедр підкидають і дивляться якою гранню він впаде на горизонтальну площину.

Нехай A_1={на грані, що випала, є синій колір}, A_2={на грані, що випала, є червоний колір}, A_3={на грані, що випала, є зелений колір}. Тетраедр однорідний, тому всі грані випадають з однаковими ймовірностями. Очевидно, що $\mathbf{P}(A_i) = \frac{1}{2}$, $\mathbf{P}(A_i \cdot A_j) = \frac{1}{4}$ при $i \neq j$. Тому події A_1, A_2, A_3 попарно незалежні. Але

$$\mathbf{P}(A_1 \cdot A_2 \cdot A_3) = \frac{1}{4} \neq \frac{1}{8} = \mathbf{P}(A_1) \cdot \mathbf{P}(A_2) \cdot \mathbf{P}(A_3),$$

тому ці події незалежними в сукупності не є.

1.4 Формули повної ймовірності та Байєса

Розглянемо деякий ймовірнісний простір $(\Omega, \mathcal{F}, \mathbf{P})$.

Означення. Набір подій H_1, $H_2, \ldots$, H_n будемо називати *повною групою подій*, якщо для них виконується таке:

1. $H_i \cdot H_j = \emptyset$, $i \neq j$;
2. $\bigcup\limits_{i=1}^{n} H_i = \Omega$.

Теорема 1. Нехай A — деяка подія, H_1, $H_2, \ldots$, H_n — повна група подій і $\mathbf{P}(H_i) \neq 0$ при всіх $i = \overline{1, n}$. Тоді має місце *формула повної ймовірності*

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

Доведення. Очевидно, що подію A можна подати у вигляді $A = \bigcup\limits_{i=1}^{n} A \cdot H_i$ і $(A \cdot H_i) \cdot (A \cdot H_j) = \emptyset$ при $i \neq j$. Тому

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

За теоремою множення ймовірностей

$$\mathbf{P}(A \cdot H_i) = \mathbf{P}(A/H_i) \cdot \mathbf{P}(H_i)$$

при всіх $i = \overline{1, n}$. Отже

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

В багатьох випадках система гіпотез не визначається однозначно. В таких випадках перевагу потрібно віддавати тій системі, для якої умовні ймовірності визначаються найбільш просто.

Приклад 1. З колоди 52 карт випадково послідовно і без повернення вибирають дві карти. Яка ймовірність того, що другою картою можна побити першу? (Друга карта повинна бути більш старшою тієї ж масті.)

Нехай A — подія, ймовірність якої нас цікавить. Розглянемо такі гіпотези: H_k ={в складі двох витягнутих карт є рівно k картинок}, $k = 0, 1, 2$ ("картинки— валет, дама, король і туз).

Очевидно, що набір H_k утворює повну групу подій. Але обчислення умовних ймовірностей $\mathbf{P}(A/H_k)$ є не менш складним ніж обчислення ймовірності події A. Це пояснюється тим, що зв'язок події A з даними гіпотезами не може бути достатньо просто описаний на мові операцій над подіями.

Розглянемо ще один набір, що утворює повну групу подій: H_k ={перша витягнута карта має вартість k очок}, $k = 2, 3, \ldots, 14$ ($k = 2$ — двійка, $k = 3$ — трійка,..., $k = 11$ — валет, $k = 12$ — дама, $k = 13$ — король, $k = 14$ — туз). Обчислимо умовні ймовірності.

$$\mathbf{P}(A/H_k) = \mathbf{P}\{\text{друга карта тієї ж масті та її вартість}\\ \text{не менша, ніж } k+1 \text{ очко}\} = \tfrac{14-k}{51}$$

за класичним означенням ймовірності.

Безумовні ймовірності гіпотез $\mathbf{P}(H_k) = \frac{1}{13}$ через рівноймовірність подій H_k.

За формулою повної ймовірності

$$\mathbf{P}(A) = \sum_{k=2}^{14} \mathbf{P}(A/H_k)\mathbf{P}(H_k) = \frac{1}{13}\sum_{k=2}^{14}\frac{14-k}{51} = \frac{2}{17}.$$

Можна запропонувати ще більш вдалий підбір повної групи подій: H_1 ={дві витягнуті карти однієї масті}, H_2 ={витягнуті карти різних мастей}.

Очевидно, що $\mathbf{P}(A/H_2) = 0$, а $\mathbf{P}(A/H_1) = 0.5$ через рівноймовірність подій {друга карта старша від першої}, {перша карта старша від другої}. Крім того,

$$\mathbf{P}(H_1) = \frac{52 \cdot 12}{52 \cdot 51} = \frac{4}{17}.$$

Отже, за формулою повної ймовірності

$$\mathbf{P}(A) = \mathbf{P}(A/H_1) \cdot \mathbf{P}(H_1) = \frac{1}{2} \cdot \frac{4}{17} = \frac{2}{17}.$$

Теорема 2. Нехай подія A така, що $\mathbf{P}(A) \neq 0$ і $H_1, H_2, \ldots, H_n$ — повна група подій, для яких $\mathbf{P}(H_i) \neq 0$ при всіх $i = \overline{1,n}$. Тоді (*формули Байєса*)

$$\mathbf{P}(H_i/A) = \frac{\mathbf{P}(A/H_i) \cdot \mathbf{P}(H_i)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A/H_i) \cdot \mathbf{P}(H_i)}{\sum\limits_{i=1}^{n} \mathbf{P}(A/H_i) \cdot \mathbf{P}(H_i)}$$

при $i = \overline{1,n}$.

Доведення. Зауважимо, що друга рівність в твердженні теореми є наслідком формули повної ймовірності.

Доведемо першу рівність. Для цього розглянемо при $i = \overline{1,n}$

$$\mathbf{P}(A \cdot H_i) = \mathbf{P}(A/H_i) \cdot \mathbf{P}(H_i) = \mathbf{P}(H_i/A) \cdot \mathbf{P}(A).$$

Звідси $\mathbf{P}(H_i/A) = \frac{\mathbf{P}(A/H_i) \cdot \mathbf{P}(H_i)}{\mathbf{P}(A)}$.

Зауваження.

1. Елементи повної групи подій ще називаються гіпотезами.

2. В формулах повної ймовірності та Байєса замість повної групи подій можна взяти набір подій (гіпотез) $H_1, H_2, \ldots, H_k$ таких, що

$$A \subset \bigcup_{i=1}^{n} H_i \text{ і } H_i \cdot H_j = \emptyset \text{ при } i \neq j.$$

3. Ймовірності $\mathbf{P}(H_i)$ в формулах Байєса часто називають апріорними[3] (до дослідними, заздалегідь відомими), а ймовірності $\mathbf{P}(H_i/A)$ — апостеріорними[4] (знайденими після того, як дослід виконано і відбулась подія A).

Особливе значення формула Байєса має для тих експериментів, в яких гіпотези H_k безпосередньо не спостерігаються, хоча апріорні ймовірності $\mathbf{P}(H_k)$ та відповідні умовні ймовірності $\mathbf{P}(A/H_k)$ відомі з додаткових досліджень. Така ситуація може виникнути, наприклад, якщо відсутній прилад, який дозволяє реєструвати події, що є гіпотезами, або реєстрація гіпотез приводить до знищення предмету спостереження. Для таких експериментів переоцінка ймовірностей гіпотез після досліду може бути здійснена на базі події A, що спостерігається і тісно пов'язана з гіпотезами. Такий підхід часто використовується в задачах медичної та технічної діагностики.

Приклад 2. Вивчається три види дефектів запам'ятовуючих пристроїв, що виконані на інтегральних мікросхемах: дефекти схем (гіпотеза H_1), дефекти, пов'язані з паразитними зв'язками між чарунками (гіпотеза H_2), та дефекти адресних шин (гіпотеза H_3). Діагностика проводиться з допомогою деяких тестів. Нехай тест проведено і виявлено певний результат (подія A). З попередніх досліджень відомо, що одержаний результат спостерігається в 40%, 20% та 30% випадків при наявності дефектів, що складають зміст гіпотез H_1, H_2 та H_3, відповідно. Крім того, в тих же дослідженнях встановлено, що, в середньому, 10% несправних запам'ятовуючих пристроїв мають дефекти схем, 60% мають дефекти зв'язків та 30% мають дефекти адресних шин.

Встановити, яка з гіпотез має найбільшу апостеріорну ймовірність (який з дефектів найбільш ймовірний).

З умови зрозуміло, що

$$\mathbf{P}(A/H_1) = 0.4, \quad \mathbf{P}(A/H_2) = 0.2, \quad \mathbf{P}(A/H_3) = 0.3,$$
$$\mathbf{P}(H_1) = 0.1, \quad \mathbf{P}(H_1) = 0.6, \quad \mathbf{P}(H_1) = 0.3.$$

За формулою повної ймовірності

$$\mathbf{P}(A) = \sum_{k=1}^{3} \mathbf{P}(A/H_k)\mathbf{P}(H_k) = 0.25,$$

а за формулами Байєса

$$\mathbf{P}(H_1/A) = \tfrac{0.4 \cdot 0.1}{0.25} = \tfrac{4}{25}, \quad \mathbf{P}(H_2/A) = \tfrac{0.2 \cdot 0.6}{0.25} = \tfrac{12}{25},$$
$$\mathbf{P}(H_3/A) = \tfrac{0.3 \cdot 0.3}{0.25} = \tfrac{9}{25}.$$

Отже, найбільш ймовірно, що запам'ятовуючий пристрій має дефекти зв'язку.

1.5 Повторні незалежні випробування

1.5.1 Схема Бернуллі

Нехай в деякому стохастичному експерименті може наблюдатися подія A, ймовірність якої дорівнює p. Розглянемо таку схему проведення випробувань (схема Бернуллі). Згаданий стохастичний експеримент проводиться в однакових умовах незалежно один від одного n

[3] a priori — до випробування

[4] a posteriori — після випробування

раз. В кожному експерименті встановлюється, відбулась (успіх) чи не відбулась (невдача) подія A.

Розглянемо деякі питання, пов'язані із схемою Бернуллі.

Спочатку знайдемо ймовірність $P_n(k)$ того, що в описаній схемі випробувань відбулось рівно k успіхів
$(k = \overline{0,n})$. Для цього розглянемо події $A_i =${в i-тому випробуванні відбувся успіх}. Через незалежність випробувань можна вважати, що події A_i незалежні в сукупності. Нехай подія B полягає в тому, що в схемі з n випробувань відбудеться рівно k успіхів. Очевидно, що

$$B = \sum A_{i_1} \cdot A_{i_2} \ldots A_{i_k} \cdot \overline{A_{j_1}} \cdot \overline{A_{j_2}} \ldots \overline{A_{j_{n-k}}},$$

де сума береться по найможливіших наборах (невпорядкованих) індексів i_1, $i_2, \ldots$, i_k причому набір j_1, $j_2, \ldots$, j_{n-k} доповнює набір i_1, $i_2, \ldots$, i_k до невпорядкованого набору чисел 1, $2, \ldots$, n.

Всі події

$$A_{i_1} \cdot A_{i_2} \ldots A_{i_k} \cdot \overline{A_{j_1}} \cdot \overline{A_{j_2}} \ldots \overline{A_{j_{n-k}}}$$

попарно несумісні і мають ймовірності $p^k(1-p)^{n-k}$. Тому

$$P_n(k) = C_n^k p^k (1-p)^{n-k}. \tag{1.7}$$

1.5.2 Найімовірніша кількість успіхів в схемі Бернуллі

Дослідимо питання про поведінку чисел $P_n(k)$, заданих формулою (1.7), при зростанні k від 0 до n. Для зручності позначимо $q = 1 - p$ і розглянемо

$$\frac{P_n(k+1)}{P_n(k)} = \frac{C_n^{k+1} p^{k+1} q^{n-k-1}}{C_n^k p^k q^{n-k}} = \frac{n!k!(n-k)!}{(k+1)!(n-k-1)!n!} \cdot \frac{p}{q} = \\ = \frac{n-k}{k+1} \cdot \frac{p}{q}.$$

Звідси $P_n(k+1) > P_n(k)$, якщо $\frac{n-k}{k+1} \cdot \frac{p}{q} > 1$, тобто

$$(n-k)p > (k+1)q \Leftrightarrow np > k(p+q) + q = k + q \Leftrightarrow k < np - q.$$

Аналогічно $P_n(k+1) < P_n(k)$, при $k > np - q$.

Якщо $np - q$ ціле, то $P_n(k+1) = P_n(k)$ при $k = np - q$. Тут слід зауважити, що $-1 < np - q < n$ в усіх випадках.

Звідси можна зробити такий висновок щодо зміни $P_n(k)$ при рості k: якщо k змінюється від 0 до $[np-q]$, то $P_n(k)$ зростає; якщо k змінюється від $[np-q]+1$ до n, то $P_n(k)$ спадає. Тому має місце таке твердження:

Теорема 1. Якщо $np-q$ $(q = 1-p)$ ціле, то є два найбільш ймовірні значеня кількості успіхів в n випробуваннях Бернуллі: $np-q$ і $np-q+1$. Якщо ж $np-q$ не ціле, то найбільш ймовірна кількість успіхів дорівнює $[np-q]+1$.

1.5.3 Теорема Пуассона

Формула (1.7) дає можливість обчислювати ймовірності $P_n(k)$ при будь-яких значеннях параметрів схеми Бернуллі. Але, особливо при великих n, використаня (1.7) веде за собою неоправдано складні обчислення. Розглянемо один спосіб наближеного обчислення чисел $P_n(k)$.

Розглянемо послідовність серій випробувань за схемою Бернуллі. В n-тій серії виконується n випробувань, ймовірність успіху в кожному з них дорівнює $p_n = \frac{\lambda}{n}$ ($\lambda = const$, $\lambda < n$).

Має місце теорема *Пуассона.*

Теорема 2. Якщо $P_n(k)$ — ймовірність k успіхів у серії з n випробувань за схемою Бернуллі, в кожному з яких ймовірність успіху дорівнює $p_n = \frac{\lambda}{n}$, то

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

Доведення. За формулою (1.7)

$$P_n(k) = C_n^k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} =$$
$$= \frac{n(n-1)\ldots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} =$$
$$= \frac{\lambda^k}{k!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{-k}\left(1-\frac{\lambda}{n}\right)^{n}.$$

Тому

$$\lim_{n\to\infty} P_n(k) = \frac{\lambda^k}{k!}\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{n}\right)\times$$
$$\times\left(1-\frac{\lambda}{n}\right)^{-k}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} =$$
$$= \frac{\lambda^k}{k!}\lim_{n\to\infty}\left(\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right)^{-\lambda} = \frac{\lambda^k}{k!}e^{-\lambda}.$$

З теореми Пуассона випливає, що в схемі Бернуллі з досить великою кількістю випробувань n та ймовірністю успіху в кожному випробуванні p

$$P_n(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}, \tag{1.8}$$

де $\lambda = np$.

Формула (1.8) дає досить точне наближення при малих p (менших 0.1) і досить великих n (не менших кількох десятків) та $np \le 10$. Хоча і ці побажання не можна вважати обов'язковими.

Приклад 1. Апаратура містить 2000 однаково надійних елементів, ймовірність відмови кожного з них дорівнює $p = 0.0005$. Знайти ймовірність відмови апаратури, якщо вона настає при відмові хоча б одного з елементів.

Оскільки кожен з елементів виходить з ладу зовсім незалежно від інших, то ми маємо справу із схемою Бернуллі з $n = 2000$ випробувань та ймовірністю успіху (вихід з ладу) в кожному випробуванні $p = 0.0005$. Тому ймовірність відмови апаратури дорівнює

$$1 - P_n(0) = 1 - C_{2000}^0(0.0005)^0(0.9995)^{2000} =$$
$$= 1-(0.9995)^{2000} \approx 0.63227$$

Скористаємось теоремою Пуассона $P_n(0) \approx \frac{\lambda^0}{0!}e^{-\lambda}$, де $\lambda = np = 1$

Тому шукана ймовірність

$$1 - P_n(0) \approx 1 - e^{-1} \approx 0.63212.$$

Як бачимо, значення, одержане з допомогою наближеної формули за теоремою Пуассона, досить добре наближає точне значення ймовірності.

Вправи до розділу 1

0.1. Монету підкидають три рази. Описати простір елементарних подій. Описати події: $A = \{$ один раз з'явиться герб$\}$, $B = \{$ при другому підкиданні з'явиться герб$\}$. Описати події: $A \cap B$, $A \cup \overline{B}$, $\overline{A}$.
0.2. Випадковий експеримент полягає в тому, що вимірюють дві величини ξ_1 і ξ_2, які можуть приймати всі значення з відрізка $[0; 1]$. Описати простір елементарних подій. Описати події: $A = \{$ максимальне із значень ξ_1 і ξ_2 — менше a $\}$; $B = \{$ вектор $(\xi_1; \xi_2)$ має довжину, яка не перевищує b $\}$; $C = \{$ величина, обернена до ξ_1 не перевищує c $\}$. Описати події: $A \cap B$, $A \cap C$, $B \cap C$, $A \cap B \cap C$.
0.3. Із колоди з 52 карт витягують навмання 4 карти. Знайти ймовірності наступних подій: $A = \{$в одержаній вибірці всі карти бубнової масті$\}$, $B = \{$вибрали хоча б один туз$\}$.
0.4. Гральний кубик підкидають 6 разів. Знайти ймовірність того, що: а) випадуть всі шість граней; б) випадуть принаймі дві однакові грані; в) випадуть тільки три різні грані.
0.5. Яка ймовірність того, що сума трьох випадково взятих відрізків, довжина кожного з яких не перевищує l, буде більшою l?
0.6. На відрізку довжиною l навмання вибираються дві точки. Знайти ймовірність того, що з одержаних відрізків можна скласти трикутник.
0.7. При переливанні крові потрібно враховувати групу крові донора і хворого. Людині, що має четверту групу крові, можна перелити кров будь-якої групи; людині з другою або третьою групою крові можна перелити кров або тієї ж групи, або першої; людині з першою групою крові можна перелити тільки кров першої групи. Серед населення 33.7% мають першу, 37.5% — другу, 20,9% — третю і 7.9% — четверту групи крові. Знайти ймовірність того, що випадково взятому хворому можна перелити кров випадково взятого донора.
0.8. В ящику лежать 20 тенісних м'ячів, серед них 15 нових і 5 вживаних. Для гри випадково вибираються два м'ячі і після гри повертаються назад. Потім для другої гри також випадково вибираються ще два м'ячі. Яка ймовірність того, що друга гра буде проводитись новими м'ячами?
0.9. На вхід радіолокаційного пристрою з ймовірністю 0.8 поступає суміш корисного сигналу з шумом, а з ймовірністю 0.2 — тільки шум. Якщо поступає корисний сигнал з шумом, то пристрій реєструє наявність якогось сигналу з ймовірністю 0.7; якщо тільки шум, — то з ймовірністю 0.3. Відомо, що пристрій зареєстрував наявність якогось сигналу. Знайти ймовірність того, що в його складі є корисний сигнал.

Розділ 2

Випадкові величини

2.1 Випадкові величини, їх розподіли

2.1.1 Функція розподілу

Нехай задано деякий ймовірнісний простір $(\Omega, \mathcal{F}, \mathbf{P})$. Нагадаємо, що це означає, що певному стохастичному експерименту відповідає простір елементарних подій Ω, σ-алгебра подій $\mathcal{F}$ та ймовірнісна міра $\mathbf{P}$. Нехай в цьому експерименті вимірюється значення деякої величини ξ. Зрозуміло, що для кожної елементарної події $\omega \in \Omega$ одержуємо деяке значення ξ, тобто, ξ є функцією від ω

Означення. Функцію $\xi(\omega)$ зі значеннями в R, задану на просторі елементарних подій Ω, будемо називати *випадковою величиною*, якщо $\{\omega : \xi(\omega) < x\} \in \mathcal{F}$ при всіх $x \in R$.

Приклад 1. Двічі підкидається монета ; ξ — кількість гербів, що випадуть.

Зрозуміло, що $\Omega = \{$гг, гц, цг, цц$\}$, $\mathcal{F}$ — σ-алгебра всіх підмножин Ω, $\mathbf{P}$ така, що $\mathbf{P}(\omega_i) = \frac{1}{4}$, $i = \overline{1,4}$.

Звідси величина ξ може приймати лише значення 0, 1, 2. Тому при $x \le 0$

$$\{\omega : \xi(\omega) < x\} = \emptyset \in \mathcal{F},$$

при $0 < x \le 1$

$$\{\omega : \xi(\omega) < x\} = \{\omega : \xi(\omega) = 0\} = \{\text{цц}\} \in \mathcal{F},$$

при $1 < x \le 2$

$$\begin{gathered}\{\omega : \xi(\omega) < x\} = \{\omega : \xi(\omega) = 0 \text{ або} \\ \xi(\omega) = 1\} = \{\text{цц, цг, гц}\} \in \mathcal{F},\end{gathered}$$

при $x > 2$

$$\begin{gathered}\{\omega : \xi(\omega) < x\} = \{\omega : \xi(\omega) = 0 \text{ або} \\ \xi(\omega) = 1 \text{ або } \xi(\omega) = 2\} = \\ = \{\text{цц, гц, цг, гг}\} = \Omega \in \mathcal{F}.\end{gathered}$$

Отже, ξ є випадковою величиною.

Зауваження. В більшості випадків при позначенні випадкової величини аргумент її, як функції на Ω, будемо пропускати і вважатимемо, що ξ і $\xi(\omega)$ це одне і те ж.

Означення. Функція $F(x)$, задана на R, така, що

$$F(x) = \mathbf{P}(\xi < x),$$

називається *функцією розподілу* випадкової величини ξ.

Приклад 2. Для випадкової величини ξ , розглянутої в прикладі 1.

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{4}, & 0 < x \le 1 \\ \frac{3}{4}, & 1 < x \le 2 \\ 1, & x > 2 \end{cases} .$$

Розглянемо основні властивості функції розподілу:

1. Функція $F(x)$ неспадна, тобто, якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$.
2. Для всіх $x \in R$ $0 \le F(x) \le 1$ і $\lim\limits_{x\to-\infty} F(x) = 0$, а $\lim\limits_{x\to+\infty} F(x) = 1$.
3. Функція $F(x)$ неперервна зліва.

Доведення. 1) Очевидно, що

$$\{\omega : \xi(\omega) < x_1\} \subset \{\omega : \xi(\omega) < x_2\},$$

тому $\mathbf{P}(\xi < x_1) \le \mathbf{P}(\xi < x_2)$ за властивістю монотонності ймовірності. Отже, $F(x_1) \le F(x_2)$.

2) Оскільки $0 \le \mathbf{P}(\xi < x) \le 1$ за властивістю ймовірності, то $0 \le F(x) \le 1$.

Розглянемо послідовність $\{x_n : n \ge 1\}$ таку, що $x_1 > x_2 > \ldots > x_n > \ldots$ і $\lim\limits_{n\to\infty} x_n = -\infty$. Та розглянемо події $A_n = \{\omega : \xi(\omega) < x_n\}$. Очевидно, що $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ і $\bigcap\limits_{n=1}^{\infty} A_n = \emptyset$. Тому за теоремою неперервності ймовірності

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \mathbf{P}(\emptyset) = 0.$$

Отже, $\lim\limits_{x\to-\infty} F(x) = 0$.

Розглянемо тепер послідовність $\{x_n : n \ge 1\}$ таку, що $x_1 < x_2 < \ldots < x_n < \ldots$ і $\lim\limits_{n\to\infty} x_n = +\infty$. Для подій $B_n = \{\omega : \xi(\omega) \ge x_n\}$, очевидно, виконується $B_1 \supset B_2 \supset \ldots \supset B_n \supset \ldots$ і $\bigcap\limits_{n=1}^{\infty} B_n = \emptyset$. Тому за тією ж властивістю неперервності ймовірності

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mathbf{P}(\overline{B_n}) = 1 - \lim_{n\to\infty} \mathbf{P}(B_n) =$$
$$= 1 - \mathbf{P}\left(\bigcap_{n=1}^{\infty} B_n\right) = 1 - \mathbf{P}(\emptyset) = 1.$$

Отже, $\lim\limits_{x\to+\infty} F(x) = 1$.

3) Розглянемо довільну послідовність $\{x_n : n \ge 1\}$ таку, що $x_1 < x_2 < \ldots < x_n < \ldots$ і $\lim\limits_{n\to\infty} x_n = x_0$. Події $C_n = \{\omega : x_n \le \xi(\omega) < x_0\}$ такі, що

$$C_1 \supset C_2 \supset \ldots \supset C_n \supset \ldots \text{ і } \bigcap_{n=1}^{\infty} C_n = \emptyset.$$

Тому за властивістю неперервності ймовірності

$$\lim_{n\to\infty} \mathbf{P}(x_n \le \xi < x_0) = \mathbf{P}(\emptyset) = 0.$$

Тобто $\lim\limits_{n\to\infty}(\mathbf{P}(\xi < x_0) - \mathbf{P}(\xi < x_n)) = 0$ і

$\lim\limits_{n\to\infty}(F(x_0) - F(x_n)) = 0$, що еквівалентне рівності

$\lim\limits_{x\to x_0-0} F(x) = F(x_0)$.

Оскільки точка x_0 вибрана абсолютно довільно, то $F(x)$ неперервна зліва на R.

2.1.2 Ймовірність попадання випадкової величини в інтервал

Нехай випадкова величина ξ має функцію розподілу $F(x)$. В багатьох випадках виникає необхідність обчислити ймовірність попадання випадкової величини ξ в один з інтервалів $(a;b)$, $[a;b)$, $(a;b]$, $[a;b]$ і т.д.

Доведемо таке твердження .

Теорема 1. Для випадкової величини ξ з функцією розподілу $F(x)$ та чисел $a, b \in R$ виконується

1. $\mathbf{P}(\xi \in (-\infty;\ b)) = F(b)$;
2. $\mathbf{P}(\xi \in (-\infty;\ b]) = F(b+0)$;
3. $\mathbf{P}(\xi \in (a;\ +\infty)) = 1 - F(a+0)$;
4. $\mathbf{P}(\xi \in [a;\ +\infty)) = 1 - F(a)$;
5. $\mathbf{P}(\xi \in (a;\ b)) = F(b) - F(a+0)$;
6. $\mathbf{P}(\xi \in (a;\ b]) = F(b+0) - F(a+0)$;
7. $\mathbf{P}(\xi \in [a;\ b)) = F(b) - F(a)$;
8. $\mathbf{P}(\xi \in [a;\ b]) = F(b+0) - F(a)$.

Доведення. Перше твердження не потребує доведення, бо співпадає, посуті, з означенням $F(b)$.

Нехай $\{x_n : n \geq 1\}$ довільна послідовність така, що $x_1 > x_2 > \ldots > x_n > \ldots$ і $\lim\limits_{n\to\infty} x_n = b$. Для подій $A_n = \{\xi < x_n\}$ виконується $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ і $\bigcap\limits_{n=1}^{\infty} A_n = \{\xi \leq b\}$. Тому за теоремою неперервності ймовірності $\lim\limits_{n\to\infty} \mathbf{P}(\xi < x_n) = \mathbf{P}(\xi \leq b)$, тобто,

$$\mathbf{P}(\xi \in (-\infty;\ b]) = \mathbf{P}(\xi \leq b) = \lim_{n\to\infty} F(x_n) = F(b+0).$$

Отже, друге твердження доведено.

Оскільки $\{\xi \in (a;\ +\infty)\} = \overline{\{\xi \in (-\infty;\ a]\}}$, то з 2) маємо $\mathbf{P}(\xi \in (a;\ +\infty)) = 1 - F(a+0)$, що є змістом твердження 3).

Так само, оскільки $\{\xi \in [a;\ +\infty)\} = \overline{\{\xi \in (-\infty;\ a)\}}$, то з 1) маємо $\mathbf{P}(\xi \in [a;\ +\infty)) = 1 - F(a)$, як і стверджувалось в 4).

Подамо кожну з подій пунктів 5) — 8) у вигляді різниці

$$\{\xi \in (a;\ b)\} = \{\xi \in (-\infty;\ b)\} \setminus \{\xi \in (-\infty;\ a]\},$$

$$\{\xi \in (a;\ b]\} = \{\xi \in (-\infty;\ b]\} \setminus \{\xi \in (-\infty;\ a]\},$$

$$\{\xi \in [a;\ b)\} = \{\xi \in (-\infty;\ b)\} \setminus \{\xi \in (-\infty;\ a)\},$$

$$\{\xi \in [a;\ b]\} = \{\xi \in (-\infty;\ b]\} \setminus \{\xi \in (-\infty;\ a)\}.$$

Оскільки кожна перша подія в різниці є наслідком другої, то з 1) і 2) випливають твердження 5) — 8).

Зауваження. Якщо функція розподілу $F(x)$ випадкової величини ξ неперервна, то, очевидно, всі ймовірності в 5) — 8) теореми дорівнюють $F(b) - F(a)$. Так само ймовірності в 1), 2) дорівнюють $F(b)$ а в 3) і 4) — $1 - F(a)$.

2.1.3 Неперервні та дискретні випадкові величини

Як ми бачили з попереднього параграфу, властивості випадкової величини багато в чому залежать від характеру її функції розподілу.

Означення. Випадкова величина ξ називається *неперервною*, якщо її функція розподілу $F(x)$ може бути подана у вигляді

$$F(x) = \int\limits_{-\infty}^{x} f(y)\mathrm{d}y \tag{2.1}$$

При цьому функція $f(x)$ називається *щільністю розподілу* величини ξ.

Зауваження.

1. З математичного аналізу відомо, що функція задана рівністю (2.1) неперервна, а у випадку неперервності функції $f(x)$ ще й диференційовна. Тому в точках, де існує похідна функції $F(x)$, $f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x}$.
2. Іноді функцію $F(x)$ називають інтегральною, а функцію $f(x)$ — диференціальною функцією неперервної випадкової величини ξ.
3. Для неперервної випадкової величини ξ ймовірності з пунктів 5) — 8) теореми 1 дорівнюють $\int\limits_a^b f(x)\mathrm{d}x$.

Наведемо найпростіші властивості щільності розподілу $f(x)$, що випливають з відповідних властивостей функції розподілу.

1. Функція $f(x)$ інтегровна в невласному розумінні на $(-\infty;\ +\infty)$ і

$$\int\limits_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1 \tag{2.2}$$

2. Функція $f(x)$ невід'ємна.

Доведення. Властивість 1) випливає з того, що

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

і має місце (2.1).

Для доведення 2) припустимо, що на деякому інтервалі $(\alpha;\ \beta)$ $f(x) < 0$. Тоді $\mathbf{P}(\xi \in (\alpha;\ \beta) = \int\limits_\alpha^\beta f(x)\mathrm{d}x < 0$, що неможливо. Тому $f(x) \geq 0$.

Зауважимо, що в окремих точках (якщо їх зліченна кількість) $f(x)$ може бути від'ємною. Хоча ми будемо такі функції змінювати так, щоб $f(x)$ було невід'ємним при всіх значеннях аргумента. При цьому, значення інтеграла в (2.1) не змінюється.

Таким чином, завжди будемо вважати, що щільність розподілу $f(x)$ невід'ємна і для неї виконується (2.2). І навпаки, кожна невід'ємна функція, для якої виконується (2.2), може бути щільністю розподілу.

Означення. Випадкова величина називається *дискретною*, якщо вона може приймати тільки зліченну кількість значень.

Нехай дискретна випадкова величина ξ приймає значення $x_1,\ x_2,\ \ldots$ і $\mathbf{P}(\xi = x_k) = p_k$. Набір пар чисел $(x_1;\ p_1),\ (x_2;\ p_2), \ldots$ будемо називати розподілом дискретної випадкової величини ξ. Ми будемо відкидати ті значення x_k випадкової величини ξ, для яких $p_k = 0$. Тому в розподілі дискретної випадкової величини всі $p_k > 0$ і, очевидно, $\sum\limits_k p_k = 1$.

Функція розподілу дискретної випадкової величини ξ задається рівністю

$$F(x) = \sum_{k: x_k < x} p_k$$

(порівняйте з прикладом 2).

Зауваження.

1. Вище було встановлено, що функція розподілу неперервної випадкової величини неперервна. Виявляється, що зовсім не обов'язково випадкова величина з неперервною функцією розподілу є неперервною. Такі величини називаються сингулярними.

2. Надалі ми будемо розглядати тільки неперервні та дискретні випадкові величини.

2.1.4 Розподіл функцій від випадкових величин

Нехай випадкова величина ξ має функцію розподілу $F_\xi(x)$ чи щільність розподілу $f_\xi(x)$. Розглянемо функцію $g(x)$, $x \in R$ таку, що $\eta = g(\xi)$ є випадковою величиною. Нагадаємо, що для цього необхідно і досить щоб

$$\{\omega : \eta(\omega) < x\} \in \mathcal{F} \text{ при всіх } x \in R.$$

Якщо випадкова величина ξ дискретна і має розподіл $(x_i; p_i)$, то випадкова величина $\eta = g(\xi)$ має розподіл $(y_k; q_k)$, де $y_k = g(x_{i_k})$, а $q_k = \sum\limits_{i: g(x_i) = y_k} p_i$.

Нехай тепер ξ — неперервна випадкова величина. Припустимо, для простоти, що функція $g(x)$ монотонно зростаюча. Тоді, оскільки

$$\mathbf{P}(\eta < x) = \mathbf{P}(g(\xi) < x) = \mathbf{P}(\xi < g^{-1}(x)), \text{ то}$$

$$F_\eta(x) = F_\xi(g^{-1}(x)). \tag{2.3}$$

Якщо, крім того, функція $g(x)$ неперервно диференційовна ($g'(x) \neq 0$), то випадкова величина η неперервна та її щільність розподілу

$$f_\eta(x) = \frac{\mathrm{d}F_\xi(g^{-1}(x))}{\mathrm{d}x} = f_\xi(g^{-1}(x)) \cdot \frac{1}{g'(g^{-1}(x))}. \tag{2.4}$$

Зауваження. Якщо ж функція $g(x)$ не задовольняє перераховані умови (доречі досить обмежливі), то, зрозуміло, користуватися формулами (2.3) чи (2.4) не можна. Тому в конкретних задачах більш прийнятним є проведення всіх викладок, які були використані при виведенні згаданих формул.

Приклад 3. Нехай випадкова величина ξ задана щільністю розподілу $f_\xi(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$, $x \in R$. Знайти розподіл випадкової величини $\eta = \xi^2$.

Знайдемо

$$F_\eta(x) = \mathbf{P}(\eta < x) = \mathbf{P}(\xi^2 < x) =$$
$$= \begin{cases} 0, & x \leq 0 \\ \mathbf{P}(-\sqrt{x} < \xi < \sqrt{x}), & x > 0 \end{cases}.$$

Нехай $x > 0$ і розглянемо

$$\mathbf{P}(-\sqrt{x} < \xi < \sqrt{x}) = \frac{1}{\pi} \int\limits_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{1+y^2} \mathrm{d}y =$$

$$= \frac{1}{\pi} \operatorname{arctg} y \Big|_{-\sqrt{x}}^{\sqrt{x}} = \frac{1}{\pi} \operatorname{arctg} \sqrt{x}.$$

Звідси

$$f_\eta(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{\pi} \cdot \frac{1}{(1+x)\sqrt{x}} & x > 0 \end{cases}.$$

2.2 Випадкові вектори

2.2.1 Розподіли випадкового вектора та його елементів

Почнемо з означення:

Означення. Впорядкований набір випадкових величин $(\xi_1; \ldots; \xi_n)$, заданих на одному й тому ж ймовірнісному просторі $(\Omega,\ \mathcal{F},\ \mathbf{P})$, будемо називати n-вимірним *випадковим вектором.*

Надалі для спрощення записів розглядатимемо тільки двовимірні випадкові вектори.

Означення. Функція двох змінних $F(x, y)$, називається *функцією розподілу* випадкового вектора $(\xi; \eta)$, якщо для всіх $x,\ y \in R$

$$F(x, y) = \mathbf{P}(\xi < x,\ \eta < y).$$

Означення. Фунція двох змінних $f(x, y)$, називається *щільністю розподілу* випадкового вектора $(\xi; \eta)$, якщо його функція розподілу може бути подана у вигляді

$$F(x, y) = \int\limits_{-\infty}^{x} \mathrm{d}s \int\limits_{-\infty}^{y} f(s, t) \mathrm{d}t.$$

В цьому випадку випадковий вектор $(\xi; \eta)$ називається *неперервним.*

Означення. Випадковий вектор $(\xi; \eta)$ називається *дискретним*, якщо множина його значень не більше ніж злічення.

Нехай дискретний випадковий вектор $(\xi; \eta)$ приймає значення $(x_i; y_j)$. Розглянемо набір чисел

$$p_{ij} = \mathbf{P}(\xi = x_i, \eta = y_j).$$

Набір трійок чисел $(x_i, y_j; p_{ij})$ будемо називати *розподілом* дискретного випадкового вектора $(\xi; \eta)$. Аналогічно до того, що ми мали для випадкових величин сформулюємо характерні властивості функцій та щільностей розподілів, а також розподілів (в дискретному випадку) випадкових векторів

1. (a) $\lim\limits_{x \to +\infty} F(x, y) = F_\eta(y)$, $\lim\limits_{y \to +\infty} F(x, y) = F_\xi(x)$, $\lim\limits_{x \to +\infty} \lim\limits_{y \to +\infty} F(x, y) = 1$;

   (b) $\int\limits_{-\infty}^{+\infty} f(x, y) \mathrm{d}x = f_\eta(y)$, $\int\limits_{-\infty}^{+\infty} f(x, y) \mathrm{d}y = f_\xi(x)$, $\int\limits_{-\infty}^{+\infty} \mathrm{d}x \int\limits_{-\infty}^{+\infty} f(x, y) \mathrm{d}y = 1$;

   (c) $\sum\limits_i p_{ij} = p_{\cdot j}$, $\sum\limits_j p_{ij} = p_{i\cdot}$, $\sum\limits_i \sum\limits_j p_{ij} = 1$.

   Тут $F_\xi(x)$, $F_\eta(y)$ — функції розподілу; $f_\xi(x)$, $f_\eta(y)$ — щільності розподілу; $(x_i; p_{i\cdot})$, $(y_j; p_{\cdot j})$ — розподіли випадкових величин ξ і η, відповідно.

2. Для всіх $x, y \in R$ $0 \leq F(x,y) \leq 1$, $f(x,y) \geq 0$, а також для всіх i, j $0 \leq p_{ij} \leq 1$.

3. Функція $F(x,y)$ неспадна і неперервна зліва по кожній із змінних.

4. $\mathbf{P}((\xi;\eta) \in G) = \int\int\limits_G f(x,y)\mathrm{d}x\mathrm{d}y$.

5. Для неперервного випадкового вектора в точках двічідиференційовності функції розподілу $F(x,y)$
$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}.$$

Доведемо деякі з цих властивостей.

Доведення. Для послідовності $x_1 < x_2 < \ldots < x_n < \ldots$ такої, що $\lim\limits_{n\to\infty} x_n = +\infty$ розглянемо події
$$A_n = \{\xi \geq x_n, \eta < y\}.$$
$$A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots \text{ і } \bigcap_{n=1}^{\infty} A_n = \emptyset.$$
Тому $\lim\limits_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}(\emptyset) = 0$.

Отже,
$$\begin{gathered}\lim_{n\to\infty} F(x_n, y) = \lim_{n\to\infty} \mathbf{P}(\xi < x_n, \eta < y) = \\ = \lim_{n\to\infty} \mathbf{P}(\{\eta < y\} \setminus \{\xi \geq x_n, \eta < y\}) = \\ = \lim_{n\to\infty} (\mathbf{P}(\eta < y) - \mathbf{P}(\xi \geq x_n, \eta < y)) = \\ = \mathbf{P}(\eta < y) = F_\eta(y),\end{gathered}$$
тому $\lim\limits_{x\to+\infty} F(x,y) = F_\eta(y)$.

Аналогічно доводиться, що $\lim\limits_{y\to+\infty} F(x,y) = F_\xi(x)$, і тому $\lim\limits_{x\to+\infty}\lim\limits_{y\to+\infty} F(x,y) = 1$.

Якщо випадковий вектор $(\xi;\eta)$ нерерервний, то, продиференціювавши по y рівність (тільки-що доведену)
$$\int\limits_{-\infty}^{+\infty} \mathrm{d}s \int\limits_{-\infty}^{y} f(s,t)\mathrm{d}t = F_\eta(y),$$
одержимо $f_\eta(y) = \int\limits_{-\infty}^{+\infty} f(s,y)\mathrm{d}s$.

Аналогічно $f_\xi(x) = \int\limits_{-\infty}^{+\infty} f(x,t)\mathrm{d}t$.

Розглянемо події $A_{ij} = \{\xi = x_i, \eta = y_j\}$, що утворюють повну групу подій. Оскільки $\bigcup\limits_i A_{ij} = \{\eta = y_j\}$, а $\bigcup\limits_j A_{ij} = \{\xi = x_i\}$, то
$$\sum_i p_{ij} = p_{\cdot j} = \mathbf{P}(\eta = y_j),\ \sum_j p_{ij} = p_{i\cdot} = \mathbf{P}(\xi = x_i).$$
Крім того, $\bigcup\limits_{ij} A_{ij} = \Omega$ і тому $\sum\limits_{ij} p_{ij} = 1$. Отже, 1) доведено.

Властивості 2) і 3) легко випливають з означень.

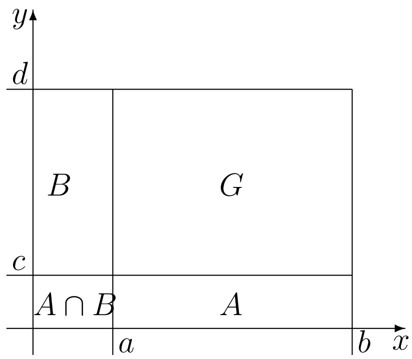


Рис. 2.1:

Нехай G — прямокутник ($G = [a; b] \times [c; d]$). Тоді (див. рис. 2.1)

$$\begin{gathered}\mathbf{P}((\xi;\eta) \in G) = \mathbf{P}((\xi;\eta) \in A \cup B \cup G) - \\ -\mathbf{P}((\xi;\eta) \in A \cup B) = \\ = \mathbf{P}((\xi;\eta) \in A \cup B \cup G) - (\mathbf{P}((\xi;\eta) \in A) + \\ +\mathbf{P}((\xi;\eta) \in B) - \mathbf{P}((\xi;\eta) \in A \cap B)).\end{gathered}$$

або

$$\begin{gathered}\mathbf{P}((\xi;\eta) \in G) = \mathbf{P}(a \le \xi \le b, c \le \eta \le d) = \\ = \mathbf{P}(\xi \le b, \eta \le d) - \mathbf{P}(\xi \le b, \eta < c) - \\ -\mathbf{P}(\xi < a, \eta \le d) + \mathbf{P}(\xi < a, \eta < c).\end{gathered}$$

З врахуванням неперервності $F(x, y)$ за обома змінними можна записати

$$\begin{gathered}\mathbf{P}((\xi;\eta) \in G) = F(b,d) - F(b,c) - F(a,d) + F(a,c) = \\ = (F(b,d) - F(b,c)) - (F(a,d) - F(a,c)) = \\ = \int\limits_a^b \mathrm{d}x \int\limits_c^d f(x,y)\mathrm{d}y.\end{gathered}$$

Якщо ж G не є прямокутником, то, використовуючи метод побудови подвійного інтегралу та неперервність ймовірнісної міри, можна встановити справедливість властивості 4).

Остання властивість легко випливає з означення 2.2.1. Дійсно $\frac{\partial F(x,y)}{\partial x} = \int\limits_{-\infty}^{y} f(x,t)\mathrm{d}t$ і $\frac{\partial^2 F(x,y)}{\partial x^2} = f(x,y)$.

Зауваження.

1. Розподіли елементів випадкового вектора іноді називають *маргінальним*. Як і для випадкових величин розподіли випадкових векторів не обмежується неперервними та дискретними розподілами, але тільки їх ми будемо розглядати.

Крім маргінальних розподілів випадкового вектора, часто приходиться розглядати і умовні розподіли. Розглянемо розподіл елемента, наприклад, ξ випадкового вектора $(\xi;\eta)$ при умові, що інший елемент цього вектора $\eta < y_0$ (у випадку неперервного випадкового вектора), чи $\eta = y_0$ (у випадку дискретного випадкового вектора). Тут y_0 — деяке фіксоване значеня, вважаємо, що $\mathbf{P}(\eta < y_0) > 0$ або $\mathbf{P}(\eta = y_0) > 0$. Зрозуміло, що

$$\mathbf{P}(\xi < x/\eta < y_0) = \frac{\mathbf{P}(\xi < x, \eta < y_0)}{\mathbf{P}(\eta < y_0)} = \frac{F(x, y_0)}{F_\eta(y_0)}$$

чи

$$\mathbf{P}(\xi = x_i/\eta = y_0) = \frac{\mathbf{P}(\xi = x_i, \eta = y_0)}{\mathbf{P}(\eta = y_0)} = \frac{p_{ij}}{p_{\cdot j}},$$

де j таке, що $y_j = y_0$. Ці рівності і задають умовний розподіл елемента ξ випадкового вектора $(\xi;\eta)$ при умові, що $\eta < y_0$ чи $\eta = y_0$.

2.2.2 Незалежні випадкові величини

Означення. Випадкові величини ξ і η називаються *незалежними*, якщо для будь-яких $x, y \in R$ події $\{\xi < x\}$, $\{\eta < y\}$ незалежні, тобто

$$\mathbf{P}(\xi < x, \eta < y) = \mathbf{P}(\xi < x) \cdot \mathbf{P}(\eta < y).$$

Зауваження. Для дискретних випадкових величин ξ і η незалежність означає

$$\mathbf{P}(\xi = x_i, \eta = y_j) = \mathbf{P}(\xi = x_i) \cdot \mathbf{P}(\eta = y_j)$$

при всіх i та j.

Знайдемо розподіл випадкового вектора $(\xi;\eta)$, елементи якого є незалежними випадковими величинами. Нехай $F_\xi(x)$ і $F_\eta(y)$ — функції розподілів випадкових величин ξ і η, відповідно. Для функції розподілу $F(x, y)$ випадкового вектора $(\xi;\eta)$

$$F(x, y) = \mathbf{P}(\xi < x, \eta < y) = \mathbf{P}(\xi < x) \cdot \mathbf{P}(\eta < y) =$$
$$= F_\xi(x) \cdot F_\eta(y).$$

Звідси — у випадку неперервних випадкових величин ξ і η — щільність розподілу випадкового вектора $(\xi;\eta)$ має вигляд

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = f_\xi(x) \cdot f_\eta(y),$$

де $f_\xi(x)$, $f_\eta(y)$ — щільності розподілів випадкових величин ξ і η, відповідно.

Якщо випадкові величини ξ і η дискретні і мають розподіли (x_i, p_i) та (y_j, q_j), відповідно, то випадковий вектор $(\xi;\eta)$ має розподіл $(x_i, y_j; p_{ij})$, де $p_{ij} = p_i \cdot q_j$ при всіх можливих значеннях i та j.

2.2.3 Композиція розподілів

Нехай ξ і η незалежні випадкові величини. Знайдемо розподіл їх суми, який будемо називати *композицією розподілів*. Припустимо, що випадкові величини ξ і η неперервні і мають щільності розподілу $f_\xi(x)$, $f_\eta(x)$, відповідно. Знайдемо щільність розподілу випадкової величини $\psi = \xi + \eta$.

Відомо, що щільність розподілу випадкового вектора $(\xi;\eta)$ має вигляд $f(x, y) = f_\xi(x) \cdot f_\eta(y)$. Тому, функція розподілу випадкової величини ψ

$$F_\psi(x) = \int\int_{s+t<x} f(s, t) ds dt = \int_{-\infty}^{+\infty} ds \int_{-\infty}^{x-s} f(s, t) dt =$$
$$= \int_{-\infty}^{+\infty} ds \int_{-\infty}^{x-s} f_\xi(s) f_\eta(t) dt$$

Звідси, обчисливши похідну від $F_\psi(x)$, одержимо

$$f_\psi(x) = \int\limits_{-\infty}^{+\infty} f_\xi(s) f_\eta(x-s)ds. \tag{2.5}$$

Зауваження. Інтеграл в правій частині формули (2.5) називається згорткою інтегровних на $(-\infty;+\infty)$ функцій $f_\xi(x)$, $f_\eta(x)$.

Отже, щільність розподілу композиції розподілів є згорткою щільностей розподілів випадкових величин, що утворюють композицію.

Якщо ж випадкові величини ξ і η залежні, то, аналогічно попередньому,

$$F_\psi(x) = \int\int\limits_{s+t<x} f(s,t)dsdt = \int\limits_{-\infty}^{+\infty} ds \int\limits_{-\infty}^{x-s} f(s,t)dt$$

і $f_\psi(x) = \int\limits_{-\infty}^{+\infty} f(s, x-s)ds$.

2.3 Числові характеристики випадкових величин

Нехай задано ймовірнісний простір $(\Omega,\ \mathcal{F},\ \mathbf{P})$. Будемо вважати, що всі випадкові величини задані на цьому просторі.

2.3.1 Математичне сподівання

Означення. *Математичним сподіванням* випадкової величини ξ, називається число $\mathbf{M}\xi = \sum\limits_i x_i p_i$, якщо ξ — дискретна випадкова величина з розподілом $(x_i; p_i)$ чи $\mathbf{M}\xi = \int\limits_{-\infty}^{\infty} xf(x)dx$, якщо ξ — неперервна випадкова величина з щільністю розподілу $f(x)$.

Зауваження. Ряд чи інтеграл з означення математичного сподівання може і розбігатися. Для таких випадкових величин математичне сподівання не існує. Надалі, якщо будемо говорити про математичне сподівання випадкової величини, то вважатимемо, що воно існує.

Розглянемо основні властивості математичного сподівання.

1. Для будь-якого числа $C \in R$

$$\mathbf{M}(C \cdot \xi) = C \cdot \mathbf{M}\xi \text{ і } \mathbf{M}C = C.$$

2. Для двох випадкових величин ξ і η

$$\mathbf{M}(\xi + \eta) = \mathbf{M}\xi + \mathbf{M}\eta.$$

3. Для двох незалежних випадкових величин ξ і η

$$\mathbf{M}(\xi \cdot \eta) = \mathbf{M}\xi \cdot \mathbf{M}\eta.$$

4. Для функцій $g(x)$ і $h(x,y)$ таких, що $g(\xi)$, $h(\xi,\eta)$ — випадкові величини

$$\mathbf{M}g(\xi) = \int\limits_{-\infty}^{+\infty} g(x)f(x)dx \quad \left(= \sum_k g(x_k)p_k\right)$$

$$\mathbf{M}h(\xi,\eta)=\int\limits_{-\infty}^{+\infty}dx\int\limits_{-\infty}^{+\infty}h(x,y)f(x,y)dy$$
$$\left(=\sum_k\sum_j g(x_k,y_j)p_{kj}\right),$$

де $f(x)$ — щільність розподілу ($(x_k;p_k)$ — розподіл) випадкової величини ξ, $f(x,y)$ ($(x_k,y_j;p_{kj})$ — розподіл) випадкового вектора $(\xi;\eta)$.

5. Якщо $\mathbf{P}(\xi\geq 0)=1$, то $\mathbf{M}\xi\geq 0$.

6. Якщо $\mathbf{P}(\xi\geq 0)=1$ і $\mathbf{M}\xi=0$, то $\mathbf{P}(\xi=0)=1$.

Доведення. Доведемо спочатку властивість 4). Для простоти обмежимось дискретним випадком. Якщо не об'єднувати однакові значення $g(x_k)$ і $h(x_k,y_j)$, то випадкові величини $g(\xi)$ та $h(\xi,\eta)$ мають такі розділи $(g(x_k),p_k)$ і $(h(x_k,y_j),p_{kj})$, відповідно . Тому, очевидно,
$\mathbf{M}g(\xi)=\sum\limits_k g(x_k)\cdot p_k$ і $\mathbf{M}h(\xi,\eta)=\sum\limits_k\sum\limits_j h(x_k,y_j)\cdot p_{kj}$,
що і потрібно було довести.

Після цього доведення 1) є, наприклад, таким

$$\mathbf{M}(C\cdot\xi)=\sum_k Cx_kp_k=C\sum_k x_kp_k=C\cdot\mathbf{M}\xi \text{ і}$$
$$\mathbf{M}C=C\cdot 1=C.$$

Так само в 2)

$$\mathbf{M}(\xi+\eta)=$$
$$=\sum_k\sum_j(x_k+y_j)p_{kj}=\sum_k\sum_j x_kp_{kj}+\sum_k\sum_j y_jp_{kj}=$$
$$=\sum_k x_kp_{k\cdot}+\sum_j y_jp_{\cdot j}=\mathbf{M}\xi+\mathbf{M}\eta.$$

Тут, як і раніше, $(x_k;p_{k\cdot})$, $(y_j;p_{\cdot j})$ — розподіли випадкових величин ξ та η, відповідно .

Для доведення 3) згадаємо, що коли ξ та η незалежні, то $p_{kj}=p_{k\cdot}\cdot p_{\cdot j}$. Тому

$$\mathbf{M}(\xi\cdot\eta)=\sum_k\sum_j x_ky_jp_{kj}=$$
$$=\sum_k\sum_j x_ky_jp_{k\cdot}p_{\cdot j}=\sum_k x_kp_{k\cdot}\sum_j y_jp_{\cdot j}=\mathbf{M}\xi\cdot\mathbf{M}\eta.$$

Властивості 5) і 6) очевидні, бо всі $x_k\geq 0$ і в сумі $\sum\limits_k x_kp_k$ всі доданки невід'ємні .

Аналогічно властивості 1) — 6) можна довести і для неперервного випадку.

Зауваження. Враховуючи, що ймовірності p_k є деякою вагою подій $\{\xi=x_k\}$ і $\sum\limits_k p_k=1$, можна назвати математичне сподівання зваженим середнім значенням випадкової величини або середнім значенням випадкової величини.

2.3.2 Дисперсія та середньоквадратичне відхилення

Означення. *Дисперсією* випадкової величини ξ називається число $\mathbf{D}\xi=\mathbf{M}(\xi-\mathbf{M}\xi)^2$.

Зауважимо, що дисперсію можуть мати тільки ті випадкові величини, що мають математичне сподівання. Далі в цьому пункті будемо розглядати випадкові величини, які мають дисперсію.

З означення випливають такі найпростіші властивості дисперсії випадкової величини.

1. Для будь-якої випадкової величини ξ
$$\mathbf{D}\xi \geq 0 \text{ і } \mathbf{D}\xi = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2.$$
2. Дисперсія невипадкової (сталої) величини
$$\mathbf{D}C = 0.$$
3. Для будь-якої випадкової величини ξ
$$\mathbf{D}(C \cdot \xi) = C^2 \mathbf{D}\xi.$$
4. Для незалежних випадкових величин ξ та η
$$\mathbf{D}(\xi + \eta) = \mathbf{D}\xi + \mathbf{D}\eta.$$
5. Для випадкової величини ξ з щільністю розподілу $f(x)$ чи розподілом $(x_k; p_k)$,
$$\mathbf{D}\xi = \int\limits_{-\infty}^{+\infty} (x - \mathbf{M}\xi)^2 f(x) dx = \int\limits_{-\infty}^{+\infty} x^2 f(x) dx - (\mathbf{M}\xi)^2$$
чи
$$\mathbf{D}\xi = \sum_k (x_k - \mathbf{M}\xi)^2 p_k = \sum_k x_k^2 p_k - (\mathbf{M}\xi)^2.$$

Доведення. Властивість 1) легко випливає з властивостей 1), 2) і 5) математичного сподівання. Дійсно, оскільки $(\xi - \mathbf{M}\xi)^2 \geq 0$ з ймовірністю 1, то
$$\mathbf{D}\xi = \mathbf{M}(\xi - \mathbf{M}\xi)^2 \geq 0.$$
Крім того,
$$\begin{aligned} \mathbf{D}\xi = \mathbf{M}(\xi - \mathbf{M}\xi)^2 = \mathbf{M}(\xi^2 - 2 \cdot \xi \mathbf{M}\xi + (\mathbf{M}\xi)^2) = \\ = \mathbf{M}\xi^2 - 2 \cdot \mathbf{M}\xi \cdot \mathbf{M}\xi + (\mathbf{M}\xi)^2 = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2. \end{aligned}$$

Далі, враховуючи, що $\mathbf{M}C = C$, одержимо
$$\mathbf{D}C = \mathbf{M}(C - C)^2 = \mathbf{M}0 = 0$$
(властивість 2)). Записавши
$$\begin{aligned} \mathbf{D}(C \cdot \xi) = \mathbf{M}(C \cdot \xi - \mathbf{M}(C \cdot \xi))^2 = \\ = \mathbf{M}(C \cdot \xi - C \cdot \mathbf{M}\xi)^2 = \mathbf{M}C^2(\xi - \mathbf{M}\xi)^2 = \\ = C^2 \mathbf{M}(\xi - \mathbf{M}\xi)^2 = C^2 \mathbf{D}\xi, \end{aligned}$$
одержимо властивість 3).

Якщо випадкові величини ξ і η незалежні, то
$\mathbf{M}(\xi \cdot \eta) = \mathbf{M}\xi \cdot \mathbf{M}\eta$. Тому
$$\begin{aligned} \mathbf{D}(\xi + \eta) = \mathbf{M}(\xi + \eta)^2 - (\mathbf{M}(\xi + \eta))^2 = \\ = \mathbf{M}(\xi^2 + 2\xi\eta + \eta^2) - (\mathbf{M}\xi + \mathbf{M}\eta)^2 = \\ = \mathbf{M}\xi^2 + 2\mathbf{M}(\xi \cdot \eta) + \mathbf{M}\eta^2 - 2\mathbf{M}\xi\mathbf{M}\eta - (\mathbf{M}\eta)^2 = \\ = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2 + \mathbf{M}\eta^2 - (\mathbf{M}\eta)^2 = \mathbf{D}\xi + \mathbf{D}\eta, \end{aligned}$$

що і стверджувалось в 4).

Властивість 5) повністю випливає з властивості 4) математичного сподівання.

Означення. *Середньоквадратичним відхиленням* випадкової величини ξ називається число $\sigma(\xi) = \sqrt{\mathbf{D}\xi}$.

Зауваження.

1. Властивість 1) дисперсії забезпечує існувавння середньоквадратичного відхилення у всіх випадкових величин, що мають дисперсію.

2. Оскільки математичне сподівання є середнім значенням випадкової величини, то дисперсія є середнім квадрату відхилення випадкової величини від свого середнього значення. Отже, дисперсія, а тому і середньоквадратичне відхилення, є характеристиками розкиду (розсіювання) значень випадкової величини відносно її середнього значення.

Приклад 1. Випадкова величина ξ задана розподілом

x_k	1	2	3
p_k	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

.

Знайти $\mathbf{M}\xi$, $\mathbf{D}\xi$, $\sigma(\xi)$.

Проведемо обчислення $\mathbf{M}\xi = 1\cdot\frac{1}{4}+2\cdot\frac{1}{4}+3\cdot\frac{1}{2} = \frac{9}{4}$; $\mathbf{D}\xi = \mathbf{M}\xi^2-(\mathbf{M}\xi)^2 = 1\cdot\frac{1}{4}+4\cdot\frac{1}{4}+9\cdot\frac{1}{2}-\frac{81}{16} = \frac{11}{16}$; $\sigma(\xi) = \sqrt{\mathbf{D}\xi} = \frac{\sqrt{11}}{4}$.

2.3.3 Характеристики випадкових векторів

Як і для випадкових величин для випадкових векторів можна ввести поняття математичного сподівання.

Означення. Вектор $(\mathbf{M}\xi; \mathbf{M}\eta)$ називається *математичним сподіванням* випадкового вектора $(\xi; \eta)$.

Аналогом дисперсії випадкової величини для випадкового вектора є коваріаційна матриця.

Означення. *Коваріацією* випадкових величин ξ і η називається число $\mathrm{cov}(\xi, \eta) = \mathbf{M}(\xi - \mathbf{M}\xi)(\eta - \mathbf{M}\eta)$. *Коефіцієнтом кореляції* двох випадкових величин ξ і η називається число $r(\xi, \eta) = \frac{\mathrm{cov}(\xi,\eta)}{\sigma(\xi)\cdot\sigma(\eta)}$.

Означення. *Коваріаційною матрицею* випадкового вектора $(\xi; \eta)$ називається матриця

$$K = \begin{pmatrix} \mathbf{D}\xi & \mathrm{cov}(\xi, \eta) \\ \mathrm{cov}(\xi, \eta) & \mathbf{D}\eta \end{pmatrix}.$$

Розглянемо деякі властивості коефіцієнта кореляції (коваріації) випадкових величин

1. $\mathrm{cov}(\xi, \eta) = \mathrm{cov}(\eta, \xi)$, $r(\xi, \eta) = r(\eta, \xi)$.

2. $|\mathrm{cov}(\xi, \eta)| \le \sigma(\xi) \cdot \sigma(\eta)$, $|r(\xi, \eta)| \le 1$.

3. Випадкові величини ξ і η лінійно залежні (тобто $\eta = a \cdot \xi + b$) тоді і тільки тоді, коли $|r(\xi, \eta)| = 1$.

4. Якщо випадкові величини ξ та η незалежні, то $r(\xi, \eta) = 0$.

Доведення. Властивість 1) елементарно випливає з означення. Для доведення властивості 2) розглянемо

$$0 \le \mathbf{M}(\eta - \mathbf{M}\eta - t(\xi - \mathbf{M}\xi))^2 = t^2\mathbf{D}\xi - 2t\mathrm{cov}(\xi,\eta) + \mathbf{D}\eta.$$

Тому дискримінант одержаного квадратного тричлена $D \le 0$, тобто,

$$4\mathrm{cov}^2(\xi,\eta) - 4\mathbf{D}\xi\mathbf{D}\eta \le 0 \text{ або } \mathrm{cov}^2(\xi,\eta) - \mathbf{D}\xi\mathbf{D}\eta \le 0$$

і тому $|\mathrm{cov}(\xi,\eta)| \le \sigma(\xi)\sigma(\eta)$. Звідси

$$|r(\xi,\eta)| = \frac{|\mathrm{cov}(\xi,\eta)|}{\sigma(\xi)\cdot\sigma(\eta)} \le 1.$$

Розглянутий квадратний тричлен має корінь тоді і тільки тоді, коли $D = 0$, тобто, $|r(\xi,\eta)| = 1$. Тому

$$\mathbf{M}(\eta - \mathbf{M}\eta - a(\xi - \mathbf{M}\xi))^2 = 0$$

і, отже, з ймовірністю 1

$$\eta - \mathbf{M}\eta - a(\xi - \mathbf{M}\xi) = 0.$$

Позначивши $\mathbf{M}\eta - a\mathbf{M}\xi = b$, одержимо $\eta = a\xi + b$, що і становить зміст властивості 3).

Властивість 4) випливає з того, що

$$\mathrm{cov}(\xi,\eta) = \mathbf{M}(\xi\eta) - \mathbf{M}\xi\mathbf{M}\eta.$$

Означення. Випадкові величини ξ і η називаються некорельованими, якщо $r(\xi,\eta) = 0$ ($\mathrm{cov}(\xi,\eta) = 0$).

Зауваження.

1. Враховуючи, що $\mathbf{D}\xi = \mathrm{cov}(\xi,\xi)$ і так само $\mathbf{D}\eta = \mathrm{cov}(\eta,\eta)$, коваріаційну матрицю випадкового вектора $(\xi;\eta)$ можна записати у вигляді

$$K = \begin{pmatrix} \mathrm{cov}(\xi,\xi) & \mathrm{cov}(\xi,\eta) \\ \mathrm{cov}(\eta,\xi) & \mathrm{cov}(\eta,\eta) \end{pmatrix},$$

   що оправдує її назву.

2. Якщо ж $|r(\xi,\eta)|$ близький до 1, то кажуть про майже лінійну залежність випадкових величин ξ і η, а коли близький до нуля, то говорять про слабку залежність (близькість до некорельованості) цих випадкових величин.

3. Взагалі кажучи, із некорельованості двох випадкових величин не випливає їх незалежність. Це підтверджує такий приклад.

Приклад 2. Нехай випадкова величина ξ задана щільністю розподілу

$$f(x) = \begin{cases} \frac{1}{2}, & \text{якщо } x \in [-1;1] \\ 0, & \text{якщо } x \notin [-1;1] \end{cases},$$

а випадкова величина $\eta = \xi^2$. Тоді

$$\mathrm{cov}(\xi,\eta) = \mathbf{M}(\xi\eta) - \mathbf{M}\xi\mathbf{M}\eta = \mathbf{M}\xi^3 - \mathbf{M}\xi\mathbf{M}\xi^2 =$$
$$= \frac{1}{2}\int_{-1}^{1} x^3dx - \frac{1}{2}\int_{-1}^{1} xdx = 0$$

2.3.4 Інші характеристики випадкових величин

Дамо означення деяких інших характеристик випадкової величини, що дещо відрізняються від розглянутих вище.

Означення. *Модою* випадкової величини ξ називається число d, що є точкою максимуму щільності розподілу випадкової величини ξ, якщо вона неперервна, чи є найбільшймовірним можливим значенням випадкової величини ξ, якщо вона дискретна.

Мода випадкової величини може бути єдиною (унімодальний розподіл) чи мати кілька значень (мультимодальний розподіл).

Означення. *Медіаною* випадкової величини ξ з функцією розподілу $F(x)$ називається значення h, для якого $F(h) \leq \frac{1}{2}$, $F(h+0) \geq \frac{1}{2}$.

Для неперервної випадкової величини медіана може бути визначена як таке значення h, для якого $F(h) = \frac{1}{2}$. Зауважимо, що медіана є значенням, яке "ділить"розподіл на дві рівні (по "вазі") частини і тому є аналогом математичного сподівання для тих розподілів, для яких математичне сподівання не існує.

Означення. *Квантилем* порядку α, $\alpha \in (0;1)$ випадкової величини (розподілу) з функцією розподілу $F(x)$ називається число p_α, для якого $F(p_\alpha) \leq \alpha$, $F(p_\alpha + 0) \geq \alpha$.

Очевидно, що медіана є квантилем порядку $\frac{1}{2}$. Якщо ж випадкова величина неперервна, то квантиль порядку α є розв'язком рівняння $F(p_\alpha) = \alpha$.

Розглянемо на завершення ще кілька характеристик випадкової величини (моментні характеристики).

Означення. *Моментом (початковим моментом)* порядку k випадкової величини ξ називається число $\mathbf{M}\xi^k$.

Означення. *Центральним моментом* порядку k випадкової величини ξ називається число $\mathbf{M}(\xi - \mathbf{M}\xi)^k$.

Зауважимо, що математичне сподівання є початковим моментом першого порядку, а дисперсія є центральним моментом другого порядку.

З моментними характеристиками пов'язані такі характеристики, як *коефіцієнт асиметрії* $\frac{\mathbf{M}(\xi-\mathbf{M}\xi)^3}{\sqrt{(\mathbf{D}\xi)^3}}$ (характеризує асиметрію в розподілі випадкової величини відносно її математичного сподівання), *коефіцієнт ексцесу* $\frac{\mathbf{M}(\xi-\mathbf{M}\xi)^4}{(\mathbf{D}\xi)^2} - 3$ (характеризує ступінь "крутості"щільності розподілу випадкової величини).

2.4 Приклади розподілів випадкових величин

2.4.1 Біноміальний розподіл

Нехай ξ — кількість успіхів в схемі Бернуллі з n випробувань. Відомо, що

$$\mathbf{P}(\xi = k) = C_n^k p^k (1-p)^{n-k},$$

де $k = 0, 1, \dots, n$, $0 < p < 1$.

Означення. Розподіл $(x_k; p_k)$, $k = 0, 1, \dots, n$, де $x_k = k$, $p_k = C_n^k p^k (1-p)^{n-k}$ називається біноміальним з параметрами n і p, а випадкова величина з цим розподілом — біноміально розподіленою.

Переконаємось, що набір чисел p_k дійсно задає розподіл ймовірностей. Для цього досить встановити, що $p_k \geq 0$ (це очевидно) і $\sum\limits_{k=0}^{n} p_k = 1$.

Дійсно

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

Знайдемо тепер математичне сподівання та дисперсію біноміального розподілу. Для цього розглянемо набір випадкових величин X_k — індикатор успіху в k-ому випробуванні схеми Бернуллі ($k = 1, 2, \ldots, n$).

$$X_k = \begin{cases} 1, & \text{якщо в } k\text{-ому випробуванні відбувся успіх,} \\ 0, & \text{в іншому випадку.} \end{cases}$$

Всі випадкові величини X_k однаково розподілені з розподілом $(1; p)$, $(0; 1-p)$ та попарно незалежні.

Оскільки $\xi = \sum\limits_{k=0}^{n} X_k$ і

$$\begin{gathered} \mathbf{M}X_k = 1 \cdot p + 0 \cdot (1-p) = p, \\ \mathbf{D}X_k = \mathbf{M}X_k^2 - (\mathbf{M}X_k)^2 = \\ = 1 \cdot p + 0 \cdot (1-p) - p^2 = p(1-p), \end{gathered}$$

то

$$\mathbf{M}\xi = \sum_{k=0}^{n} \mathbf{M}X_k = np, \quad \mathbf{D}\xi = \sum_{k=0}^{n} \mathbf{D}X_k = np(1-p).$$

2.4.2 Розподіл Пуассона

Означення. Розподіл (x_k, p_k), $k = 0, 1, 2, \ldots$, де
$x_k = k$, $p_k = \frac{\lambda^k}{k!} e^{-\lambda}$, $\lambda > 0$, називається розподілом Пуассона з параметром λ.

З розподілом Пуассона ми вже зустрічались при вивченні схеми Бернуллі. Цей розподіл згідно теореми Пуассона може бути одержаний з біноміального розподілу шляхом граничного переходу при $n \to +\infty$.

Очевидно, що $p_k > 0$ при всіх $k \geq 0$. Крім того

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

Тому набір чисел p_k дійсно може бути розподілом ймовірностей.

Математичне сподівання розподіленої за законом Пуассона випадкової величини ξ

$$\begin{gathered} \mathbf{M}\xi = \sum\limits_{k=0}^{\infty} k \tfrac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum\limits_{k=1}^{\infty} \tfrac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \\ = \lambda e^{\lambda} \cdot e^{-\lambda} = \lambda. \end{gathered}$$

Дисперсія

$$\begin{gathered} \mathbf{D}\xi = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2 = \sum\limits_{k=0}^{\infty} k^2 \tfrac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 = \\ = \sum\limits_{k=1}^{\infty} (k - 1 + 1) \tfrac{\lambda^k}{(k-1)!} e^{-\lambda} - \lambda^2 = \\ = \sum\limits_{k=2}^{\infty} \tfrac{\lambda^k}{(k-2)!} e^{-\lambda} + \sum\limits_{k=1}^{\infty} \tfrac{\lambda^k}{(k-1)!} e^{-\lambda} - \lambda^2 = \\ = \lambda^2 e^{\lambda} \cdot e^{-\lambda} + \lambda e^{\lambda} \cdot e^{-\lambda} - \lambda^2 = \lambda. \end{gathered}$$

Таким чином параметр λ розподілу Пуассона є одночасно математичним сподіванням та дисперсією випадкової величини, що має цей розподіл.

Розглянемо одну цікаву властивість розподілу Пуассона: композиція двох розподілів Пуассона з параметрами λ_1 і λ_2 є розподілом Пуассона з параметром $\lambda_1 + \lambda_2$. Дійсно, нехай незалежні випадкові величини ξ і η мають розподіл Пуассона з параметрами λ_1 і λ_2, відповідно. За формулою повної ймовірності

$$\mathbf{P}(\xi+\eta=k) = \sum_{i=0}^{k} \mathbf{P}(\xi=i)\cdot\mathbf{P}(\eta=k-i) =$$
$$= \sum_{i=0}^{k} \frac{\lambda_1^i}{i!}e^{-\lambda_1}\frac{\lambda_2^{k-i}}{(k-i)!}e^{-\lambda_2} = \frac{1}{k!}\sum_{i=0}^{k} C_k^i \lambda_1^i \lambda_2^{k-i} e^{-(\lambda_1+\lambda_2)} =$$
$$= \frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)},$$

що і треба було довести.

2.4.3 Геометричний розподіл

Розглянемо нескінченну кількість незалежних випробувань з ймовірністю успіху в кожному випробуавнні p. Нехай ξ — кількість випробувань до першого успіху.

Тоді $\mathbf{P}(\xi = k) = p(1-p)^k$.

Означення. Розподіл $(x_k; p_k)$, $k = 0, 1, 2, \ldots$, де
$x_k = k$, $p_k = p(1-p)^k$, називається геометричним розподілом з параметром p.

Перевіримо характерні властивості розподілу

1. $p_k > 0$,

2. $\sum\limits_{k=0}^{\infty} p_k = \sum\limits_{k=0}^{\infty} p(1-p)^k = \frac{p}{1-(1-p)} = 1$, бо послідовність p_k утворює нескінченну спадну геометричну прогресію із знаменником $q = 1-p$.

Для обчислення числових характеристик випадкової величини ξ розглянемо функцію $f(x) = \sum\limits_{k=0}^{\infty} x^k = \frac{1}{1-x}$, $|x| < 1$.

Тому

$$\mathbf{M}\xi = \sum_{k=0}^{\infty} kp(1-p)^k = p(1-p)\sum_{k=1}^{\infty} k(1-p)^{k-1} =$$
$$= p(1-p)f'(1-p) = p(1-p)\frac{1}{(1-(1-p))^2} = \frac{1-p}{p}.$$

$$\mathbf{D}\xi = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2 = \sum_{k=0}^{\infty} k^2 p(1-p)^k - \frac{(1-p)^2}{p^2} =$$
$$= p(1-p)^2 \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} +$$
$$+ \sum_{k=0}^{\infty} kp(1-p)^k - \frac{(1-p)^2}{p^2} =$$
$$= p(1-p)^2 f''(1-p) + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} =$$
$$= p(1-p)^2 \frac{2}{(1-(1-p))^3} + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2}.$$

2.4.4 Рівномірний розподіл

Розглянемо неперервну випадкову величину ξ, яка може приймати будь-яке значення з відрізка $[a;b]$. Нехай ймовірність того, що ξ попаде в довільний інтервал, який повністю міститься в $[a;b]$, залежить тільки від довжини цього інтервала. Зрозуміло, що щільність розподілу цієї випадкової величини повинна мати вигляд

$$f(x) = \begin{cases} c, & \text{якщо } x \in [a;b], \\ 0, & \text{якщо } x \notin [a;b]. \end{cases}$$

З умови $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$ знайдемо параметр c:

$$\int\limits_{-\infty}^{+\infty} f(x)dx = \int\limits_{a}^{b} cdx = c(b-a),$$

тому $c = \frac{1}{b-a}$.

Означення. Випадкова величина ξ має рівномірний розподіл на відрізку $[a;b]$, якщо її щільність розподілу

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

Знайдемо числові характеристики рівномірно розподіленої випадкової величини.

Математичне сподівання

$$\mathbf{M}\xi = \int\limits_{-\infty}^{+\infty} xf(x)dx = \int\limits_{a}^{b} x\frac{1}{b-a}dx = \\ = \frac{b^2-a^2}{2} \cdot \frac{1}{b-a} = \frac{a+b}{2}.$$

Дисперсія

$$\mathbf{D}\xi = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2 = \int\limits_{-\infty}^{+\infty} x^2 f(x)dx - (\mathbf{M}\xi)^2 = \\ = \int\limits_{a}^{b} x^2\frac{1}{b-a}dx - \frac{(a+b)^2}{4} = \\ = \frac{b^3-a^3}{3} \cdot \frac{1}{b-a} - \frac{(a+b)^2}{4} = \frac{(a-b)^2}{12}.$$

На завершення розглянемо спосіб моделювання розподілів маючи рівномірний розподіл.

Нехай випадкова величина η має неперервну зростаючу функцію розподілу $F(x)$. Розглянемо випадкову величину $\xi = F(\eta)$

$$\mathbf{P}(\xi < x) = \mathbf{P}(F(\eta) < x) = \\ = \begin{cases} 0, & x \le 0 \\ \mathbf{P}(\eta < F^{-1}(x)), & 0 < x \le 1 \\ 1, & x > 1 \end{cases} = \\ = \begin{cases} 0, & x \le 0 \\ F(F^{-1}(x)), & 0 < x \le 1 \\ 1, & x > 1 \end{cases} = \begin{cases} 0, & x \le 0 \\ x, & 0 < x \le 1 \\ 1, & x > 1 \end{cases}.$$

Тому щільність розподілу випадкової величини ξ

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

тобто, ξ рівномірно розподілена на $[0;1]$.

Звідси випливає спосіб моделювання будь-якого неперервного розподілу, що задається функцією розподілу на $F(x)$. Досить взяти рівномірно розподілену на відрізку $[0;1]$ випадкову величину ξ і розглянути випадкову величину $\eta = F^{-1}(\xi)$. Вона матиме функцію розподілу $F(x)$.

2.4.5 Показниковий розподіл

Означення. Випадкова величина ξ має показниковий (експоненціальний) розподіл з параметром λ, якщо її щільність розподілу

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

Перевіримо, що функція $f(x)$ може бути щільністю розподілу:

1. $f(x) > 0$;
2. $\int\limits_{-\infty}^{+\infty} f(x)dx = \int\limits_{0}^{+\infty} \lambda e^{-\lambda x}dx = -e^{-\lambda x}\Big|_0^{+\infty} = 1.$

Математичне сподівання

$$\mathbf{M}\xi = \int\limits_{-\infty}^{+\infty} xf(x)dx = \int\limits_{0}^{+\infty} x\lambda e^{-\lambda x}dx =$$
$$= -xe^{-\lambda x}\Big|_0^{+\infty} + \int\limits_{0}^{+\infty} e^{-\lambda x}dx = -\tfrac{1}{\lambda}e^{-\lambda x}\Big|_0^{+\infty} = \tfrac{1}{\lambda}.$$

Дисперсія

$$\mathbf{D}\xi = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2 = \int\limits_{-\infty}^{+\infty} x^2 f(x)dx - (\mathbf{M}\xi)^2 =$$
$$= \int\limits_{0}^{+\infty} x^2\lambda e^{-\lambda x}dx - \tfrac{1}{\lambda^2} =$$
$$= -x^2e^{-\lambda x}\Big|_0^{+\infty} + 2\int\limits_{0}^{+\infty} x\lambda e^{-\lambda x}dx - \tfrac{1}{\lambda^2} =$$
$$= \tfrac{2}{\lambda^2} - \tfrac{1}{\lambda^2} = \tfrac{1}{\lambda^2}.$$

Розглянемо на завершення

$$\mathbf{P}(\xi \in [k;k+1]) = \int\limits_{k}^{k+1} f(x)dx = \int\limits_{k}^{k+1} \lambda e^{-\lambda x}dx =$$
$$= -e^{-\lambda x}\Big|_k^{k+1} = (1 - e^{-\lambda}) \cdot e^{-k\lambda}.$$

Звідси видно, що показниковий розподіл є неперервним аналогом геометричного розподілу з параметром $p = 1 - e^{-\lambda}$.

2.4.6 Нормальний розподіл

Нормальний роподіл відіграє виключно важливу роль в теорії ймовірностей та її застосуваннях. Важливість цього розподілу пояснюється тим, що при досить широких припущеннях розподіл суми великої кількості випадкових величин виявляється близьким до нормального розполілу. Це має місце, коли на значення деякої величини впливає велика кількість рівноправних випадкових факторів.

Означення. Випадкова величина ξ має нормальний розподіл з параметрами a та σ^2 (позначається $N(a;\sigma^2)$), якщо її щільність розподілу

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

Перш ніж доводити, що функція $f(x)$ може бути щільністю розподілу випадкової величини знайдемо (інтеграл Пуассона)

$$\mathcal{I} = \int_{-\infty}^{+\infty} e^{-t^2/2} dt.$$

Зауважимо, між іншим, що невизначений інтеграл від функції $e^{-t^2/2}$ не виражається через елементарні функції.

Розглянемо

$$\begin{aligned} \mathcal{I}^2 &= \int_{-\infty}^{+\infty} e^{-t^2/2} dt \int_{-\infty}^{+\infty} e^{-s^2/2} ds = \\ &= \left| \begin{array}{c} s = \rho\cos\phi \\ t = \rho\sin\phi \end{array} \right| = \int_0^{2\pi} d\phi \int_0^{+\infty} e^{-\rho^2/2} \rho d\rho = \\ &= 2\pi e^{-\rho^2/2} \Big|_0^{+\infty} = 2\pi. \end{aligned}$$

Тому $\mathcal{I} = \sqrt{2\pi}$.

Звідси

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)dx &= \tfrac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \exp\left\{-\tfrac{(x-a)^2}{2\sigma^2}\right\} dx = \\ &= \left| \begin{array}{c} \frac{x-a}{\sigma} = t \\ dx = \sigma dt \end{array} \right| = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2} dt = 1 \end{aligned}$$

Крім того, $f(x) > 0$ при $x \in R$, тому $f(x)$ може бути щільністю розподілу випадкової величини.

З'ясуємо суть параметрів нормального розподілу. Для цього знайдемо його математичне сподівання і дисперсію.

$$\begin{aligned} \mathbf{M}\xi &= \int_{-\infty}^{+\infty} x f(x) dx = \\ &= \tfrac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x \exp\left\{-\tfrac{(x-a)^2}{2\sigma^2}\right\} dx = \\ &= \tfrac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \tfrac{x-a}{\sigma} \exp\left\{-\tfrac{(x-a)^2}{2\sigma^2}\right\} dx + a = \\ &= -\tfrac{\sigma}{\sqrt{2\pi}} \exp\left\{-\tfrac{(x-a)^2}{2\sigma^2}\right\} \Big|_{-\infty}^{+\infty} + a = a; \end{aligned}$$

$$\mathbf{D}\xi = \int\limits_{-\infty}^{+\infty} (x - \mathbf{M}\xi)^2 f(x) dx =$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} (x-a)^2 \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} dx =$$
$$= -\frac{(x-a)\sigma}{\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}\Big|_{-\infty}^{+\infty} +$$
$$+\sigma^2 \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} dx = \sigma^2$$

Отже, параметрами нормального розподілу $N(a;\sigma^2)$ є його математичне сподівання a та дисперсія σ^2.

Для багатьох практичних задач є важливим обчислення ймовірностей попадання нормально розподіленої випадкової величини в деякий інтервал

$$\mathbf{P}(\xi \in (\alpha;\beta)) = \int\limits_{\alpha}^{\beta} f(x) dx =$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{\alpha}^{\beta} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} dx = \left| \begin{array}{c} \frac{x-a}{\sigma} = t \\ dx = \sigma dt \end{array} \right| =$$
$$= \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-t^2/2} dt = \Phi(\tfrac{\beta-a}{\sigma}) - \Phi(\tfrac{\alpha-a}{\sigma}),$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt$ (функція Лапласа).

Функція Лапласа табульована, що дає можливість обчислювати розглянуті ймовірності.

Наведемо деякі властивості функції Лапласа:

1. $\lim\limits_{x\to+\infty} \Phi(x) = \frac{1}{2}$;

2. $\Phi(-x) = -\Phi(x)$;

3. $F_0(x) = \frac{1}{2} + \Phi(x)$ є функцією розподілу $N(0;1)$.

Властивість 1) випливає з властивості 3), яка є очевидною, бо щільність розподілу $N(0;1)$

$$f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

є парною функцією і

$$F_0(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt.$$

Для доведення властивості 2) запишемо

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^{-x} e^{-t^2/2} dt = \left| \begin{array}{c} t = -u \\ dt = -du \end{array} \right| =$$
$$= -\frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-u^2/2} du = -\Phi(x).$$

Зауваження. Нормальний розподіл $N(0;1)$ ще називають стандартним нормальним розподілом. Його функцію розподілу іноді також можна зустріти табульованою. Властивість 2) для неї запишеться у вигляді

$$F_0(-x) = 1 - F_0(x).$$

На завершення розглянемо ще одну властивість нормального розподілу. Нехай ξ має нормальний розподіл $N(a;\sigma^2)$. Тоді випадкова величина $\eta = k\xi + b$ $(k \neq 0)$ також має нормальний розподіл $N(ka+b;k^2\sigma^2)$.

Дійсно, нехай для визначеності $k>0$, тоді

$$\mathbf{P}(\eta < x) = \mathbf{P}(k\xi + b < x) = \mathbf{P}(\xi < \tfrac{x-b}{k}) =$$
$$= \tfrac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\frac{x-b}{k}} \exp\left\{-\tfrac{(t-a)^2}{2\sigma^2}\right\} dt =$$
$$= \left| \begin{array}{c} \frac{t-a}{\sigma} = s \\ dt = \sigma ds \end{array} \right| = \tfrac{1}{\sqrt{2\pi}k\sigma} \int\limits_{-\infty}^{x} \exp\left\{-\tfrac{\left(\frac{s-b}{k}-a\right)^2}{2\sigma^2}\right\} ds =$$
$$= \tfrac{1}{\sqrt{2\pi}k\sigma} \int\limits_{-\infty}^{x} \exp\left\{-\tfrac{(s-(ka+b))^2}{2(\sigma k)^2}\right\} ds,$$

що є функцією розподілу $N(ka+b;k^2\sigma^2)$.

Отже, шляхом перетворення $\eta = \sigma\xi + a$ з випадкової величини ξ, що має стандартний нормальний розподіл, можна одержати випадкову величину η з розподілом $N(a;\sigma^2)$.

2.4.7 Двовимірний нормальний розподіл

Означення. Випадковий вектор $(\xi;\eta)$ називається нормально розподіленим $(N(a_1;a_2;\sigma_1^2;\sigma_2^2;\rho))$, якщо його щільність розподілу $(x,y \in R)$

$$f(x,y) = \tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$
$$\times \exp\left\{-\tfrac{1}{2(1-\rho^2)}\left(\tfrac{(x-a_1)^2}{\sigma_1^2} - \tfrac{2\rho(x-a_1)(y-a_2)}{\sigma_1\sigma_2} + \tfrac{(y-a_2)^2}{\sigma_2^2}\right)\right\},$$

якщо $\sigma_1^2\sigma_2^2(1-\rho^2) \neq 0$.

Спеціально не будемо перевіряти, що вказана функція є щільністю розподілу (хоч її невід'ємність очевидна). Знайдемо маргінальні розподіли випадкового вектора $(\xi;\eta)$.

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f(x;y)dy = \tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$
$$\times \int\limits_{-\infty}^{+\infty} \exp\left\{-\tfrac{1}{2(1-\rho^2)}\left(\tfrac{(x-a_1)^2}{\sigma_1^2} - \tfrac{2\rho(x-a_1)(y-a_2)}{\sigma_1\sigma_2} + \right.\right.$$
$$\left.\left. + \tfrac{(y-a_2)^2}{\sigma_2^2}\right)\right\} dy = \tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$
$$\times \int\limits_{-\infty}^{+\infty} \exp\left\{-\tfrac{1}{2(1-\rho^2)}\left(\rho\tfrac{x-a_1}{\sigma_1} - \tfrac{y-a_2}{\sigma_2}\right)^2 - \tfrac{(x-a_1)^2}{2\sigma_1^2}\right\} dy =$$
$$= \left| \begin{array}{c} \frac{1}{\sqrt{1-\rho^2}}\left(\rho\frac{x-a_1}{\sigma_1} - \frac{y-a_2}{\sigma_2}\right) = t \\ dy = -\sigma_2\sqrt{1-\rho^2}dt \end{array} \right| =$$
$$= \tfrac{-\sigma_2\sqrt{1-\rho^2}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int\limits_{-\infty}^{+\infty} e^{-t^2/2}dt \cdot \exp\left\{-\tfrac{(x-a_1)^2}{2\sigma_1^2}\right\} =$$
$$= \tfrac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\tfrac{(x-a_1)^2}{2\sigma_1^2}\right\}.$$

Тобто ξ має нормальний розподіл $N(a_1;\sigma_1^2)$ і, аналогічно, η має нормальний розподіл $N(a_2;\sigma_2^2)$. Звідси, між іншим, випливає, що $f(x;y)$ є дійсно щільністю розподілу випадкового вектора. Крім того, параметри двовимірного нормального розподілу мають такий зміст $a_1 = \mathbf{M}\xi$, $a_2 = \mathbf{M}\eta$, $\sigma_1^2 = \mathbf{D}\xi$, $\sigma_2^2 = \mathbf{D}\eta$.

Знайдемо коефіцієнт кореляції між випадковими величинами ξ і η.

$$\begin{gathered}\operatorname{cov}(\xi,\eta)=\mathbf{M}(\xi-\mathbf{M}\xi)(\eta-\mathbf{M}\eta)=\\ =\tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int\limits_{-\infty}^{+\infty}dx\int\limits_{-\infty}^{+\infty}(x-a_1)(y-a_2)\times\\ \times\exp\left\{-\tfrac{1}{2(1-\rho^2)}\left(\tfrac{(x-a_1)^2}{\sigma_1^2}-\tfrac{2\rho(x-a_1)(y-a_2)}{\sigma_1\sigma_2}+\tfrac{(y-a_2)^2}{\sigma_2^2}\right)\right\}dy=\end{gathered}$$

$$\begin{gathered}=\tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int\limits_{-\infty}^{+\infty}dx\int\limits_{-\infty}^{+\infty}(x-a_1)(y-a_2)\times\\ \times\exp\left\{-\tfrac{1}{2(1-\rho^2)}\left(\tfrac{x-a_1}{\sigma_1}-\rho\tfrac{y-a_2}{\sigma_2}\right)^2-\tfrac{(y-a_2)^2}{2\sigma_2^2}\right\}dy=\end{gathered}$$

$$=\left|\begin{array}{c}\tfrac{1}{\sqrt{1-\rho^2}}\left(\tfrac{x-a_1}{\sigma_1}-\rho\tfrac{y-a_2}{\sigma_2}\right)=t\\ dx=\sigma_1\sqrt{1-\rho^2}dt\end{array}\right|=$$

$$\begin{gathered}=\tfrac{1}{2\pi\sigma_2}\int\limits_{-\infty}^{+\infty}dt\int\limits_{-\infty}^{+\infty}\left(\sqrt{1-\rho^2}\sigma_1t+\rho\tfrac{\sigma_1}{\sigma_2}(y-a_2)\right)(y-a_2)\times\\ \times\exp\left\{-\tfrac{t^2}{2}-\tfrac{(y-a_2)^2}{2\sigma_2^2}\right\}dy=\\ =\tfrac{\sqrt{1-\rho^2}\sigma_1}{2\pi\sigma_2}\int\limits_{-\infty}^{+\infty}te^{-t^2/2}dt\int\limits_{-\infty}^{+\infty}(y-a_2)\exp\left\{-\tfrac{(y-a_2)^2}{2\sigma_2^2}\right\}dy+\end{gathered}$$

$$\begin{gathered}+\rho\tfrac{\sigma_1}{2\pi\sigma_2^2}\int\limits_{-\infty}^{+\infty}e^{-t^2/2}dt\int\limits_{-\infty}^{+\infty}(y-a_2)\exp\left\{-\tfrac{(y-a_2)^2}{2\sigma_2^2}\right\}dy=\\ =0+\rho\tfrac{\sigma_1}{2\pi\sigma_2^2}\sqrt{2\pi}\sigma_2^2\sqrt{2\pi}\sigma_2=\rho\sigma_1\sigma_2.\end{gathered}$$

Тому кофіцієнт кореляції

$$r(\xi,\eta)=\frac{\operatorname{cov}(\xi,\eta)}{\sigma(\xi)\sigma(\eta)}=\frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2}=\rho$$

і, отже, параметр ρ двовимірного нормального розподілу є коефіцієнтом кореляції між його компонентами ξ та η.

Як відомо, коли $\rho=0$, то випадкові величини ξ і η некорельовані. Тоді

$$\begin{gathered}f(x,y)=\tfrac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\tfrac{(x-a_1)^2}{2\sigma_1^2}-\tfrac{(y-a_2)^2}{2\sigma_2^2}\right\}=\\ =f_\xi(x)\cdot f_\eta(y),\end{gathered}$$

що свідчить про незалежність випадкових величин ξ і η. Отже, для нормально розподілених випадкових величин поняття некорельованості та незалежності тотожні.

Зауваження. Нормальний розподіл на площині з параметром $\rho=0$ називається канонічним. При цьому головний еліпс розсіювання цього розподілу

$$\tfrac{1}{2(1-\rho^2)}\left(\tfrac{(x-a_1)^2}{\sigma_1^2}-\tfrac{2\rho(x-a_1)(y-a_2)}{\sigma_1\sigma_2}+\tfrac{(y-a_2)^2}{\sigma_2^2}\right)=\lambda^2$$

має канонічне рівняння відносно змінних $x-a_1$ та $y-a_2$. Ймовірність попадання нормального випадкового вектора в основний еліпс розсіювання дорівнює $1-e^{-\lambda^2}$.

Встановимо це для канонічного нормального розподілу. Ймовірність попадання випадкового вектора з цим розподілом в головний еліпс розсіювання

$$\frac{1}{2}\left(\frac{(x-a_1)^2}{\sigma_1^2}+\frac{(y-a_2)^2}{\sigma_2^2}\right)=\lambda^2$$

дорівнює

$$\begin{aligned} p_\lambda &= \int\int_D \tfrac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\tfrac{(x-a_1)^2}{\sigma_1^2}-\tfrac{(y-a_2)^2}{\sigma_2^2}\right\}dxdy = \\ &= \left|\begin{array}{c} x=a_1+\sigma_1 r\cos\phi \\ y=a_2+\sigma_2 r\sin\phi \end{array}\right| = \\ &= \int_0^{2\pi} d\phi \int_0^{\sqrt{2}\lambda} \tfrac{1}{2\pi\sigma_1\sigma_2}e^{-r^2/2}r\sigma_1\sigma_2 dr = \\ &= 2\pi\tfrac{1}{2\pi}\left(-e^{-r^2/2}\right)\Big|_0^{\sqrt{2}\lambda} = 1-e^{-\lambda^2}, \end{aligned}$$

де $D=\left\{(x;y):\frac{1}{2}\left(\frac{(x-a_1)^2}{\sigma_1^2}+\frac{(y-a_2)^2}{\sigma_2^2}\right)\le\lambda^2\right\}$.

Зауважимо, що півосі головного еліпса розсіювання дорівнюють $\sqrt{2}\lambda\sigma_1$ і $\sqrt{2}\lambda\sigma_2$.

Обчислення ймовірностей p_λ при $\lambda=1,2,3$ дають такі результати $p_1=0.632$, $p_2=0.982$, $p_3=0.999$. Звідси видно, що в канонічному випадку нормально розподілений випадковий вектор майже достовірно попадає в еліпс з центром $(a_1;a_2)$ (центр розсіювання) і півосями $3\sqrt{2}\sigma_1$ і $3\sqrt{2}\sigma_2$.

2.5 Характеристичні функції

2.5.1 Означення на властивості

Розглянемо на
ймовірніносному простору $(\Omega,\mathcal{F},\mathbf{P})$ двовимірний випадковий вектор $(\xi;\eta)$. Задамо *комплекснозначну випадкову величину* ψ рівністю $\psi=\xi+i\eta$.

Якщо існують математичні сподівання випадкових величин ξ і η, то математичне сподівання випадкової величини ψ визначимо так: $\mathbf{M}\psi=\mathbf{M}\xi+i\mathbf{M}\eta$.

Зауважимо, що введені поняття задовольняють всі звичайні для них умови. Єдине, що розподілом комплекснозначної випадкової величини слід вважати розподіл відповідного випадкового вектора.

Означення. *Характеристичною функцією* $\phi(t)$ випадкової величини ξ називається комплекснозначна функція, задана при $t\in R$ співвідношенням

$$\phi(t)=\mathbf{M}e^{it\xi}=\mathbf{M}(\cos t\xi+i\sin t\xi).$$

Якщо випадкова величина ξ неперервна і $f(x)$ — її щільність розподілу, то

$$\phi(t)=\int_{-\infty}^{+\infty}e^{itx}f(x)dx.$$

Тобто, характеристична функція є перетворенням Фур'є щільності розподілу неперервної випадкової величини. Для дискретної випадкової величини з розподілом $(x_k;p_k)$ характеристична функція

$$\phi(t)=\sum_k e^{itx_k}p_k.$$

Розглянемо деякі *властивості характеристичних функцій*:

1. $\phi(0) = 1$, $\phi(t) \le 1$ для всіх $t \in R$;

2. $\phi(-t) = \overline{\phi(t)}$

3. Для всіх $n \in N$, будь-яких дійсних чисел $t_1, t_2, \ldots, t_n$ і будь-яких комплексних чисел $c_1, c_2, \ldots, c_n$
$$\sum_{k=1}^{n}\sum_{m=1}^{n}\phi(t_m - t_k)c_m\overline{c_k} \ge 0;$$

4. Характеристична функція суми двох незалежних випадкових величин дорівнює добутку характеристичних функцій цих величин.

5. Характеристична функція однозначно визначає розподіл випадкової величини.

6. Якщо для випадкової величини ξ існує $\mathbf{M}|\xi|^n$, то її характеристична функція n раз неперервно диференційовна і $\phi^{(n)}(0) = i^n\mathbf{M}\xi^n$;

7. Нехай $F_n(x)$ — послідовність функцій розподілу, $F(x)$ — неперервна функція розподілу, $\phi_n(t)$ і $\phi(t)$ відповідні їм характеристичні функції. Якщо
$$\lim_{n\to\infty}\phi_n(t) = \phi(t)$$
для всіх $t \in R$, то
$$\lim_{n\to\infty}\sup_x |F_n(x) - F(x)| = 0,$$
тобто $F_n(x)$ рівномірно по x збігається до $F(x)$.

Доведемо деякі з цих властивостей, обмежившись випадком неперервних випадкових величин.

Доведення. Очевидно, що $\phi(0) = \mathbf{M}e^{i0\xi} = \mathbf{M}1 = 1$. Далі
$$|\phi(t)| = \left|\mathbf{M}e^{it\xi}\right| = \left|\int\limits_{-\infty}^{+\infty} e^{itx}f(x)dx\right| \le$$
$$\le \int\limits_{-\infty}^{+\infty}\left|e^{itx}\right| f(x)dx = \int\limits_{-\infty}^{+\infty} f(x)dx = 1,$$
бо $|e^{itx}| = 1$ і $f(x) \ge 0$, що і стверджувалось в 1).

Враховуючи, що
$$\phi(-t) = \mathbf{M}e^{-it\xi} = \mathbf{M}(\cos t\xi - i\sin t\xi) =$$
$$= \mathbf{M}\cos t\xi - i\mathbf{M}\sin t\xi = \overline{\mathbf{M}\cos t\xi + i\mathbf{M}\sin t\xi} =$$
$$= \overline{\mathbf{M}e^{it\xi}} = \overline{\phi(t)},$$
одержимо властивість 2).

Розглянемо
$$\sum_{k=1}^{n}\sum_{m=1}^{n}\phi(t_m - t_k)c_m\overline{c_k} =$$
$$= \mathbf{M}\left(\sum_{k=1}^{n}\sum_{m=1}^{n} e^{i(t_m - t_k)\xi}c_m\overline{c_k}\right) =$$
$$= \mathbf{M}\left(\sum_{m=1}^{n} e^{it_m\xi}c_m\overline{\sum_{m=1}^{n} e^{it_k\xi}c_k}\right) =$$
$$= \mathbf{M}\left|\sum_{m=1}^{n} e^{it_m\xi}c_m\right|^2 \ge 0,$$

що становить зміст властивості 3).

Нехай $\phi_1(t)$, $\phi_2(t)$ — характеристичні функції незалежних випадкових величин ξ_1 і ξ_2 відповідно. Тоді характеристична функція $\phi(t)$ їх суми (властивість 4))

$$\phi(t) = \mathbf{M}e^{it(\xi_1+\xi_2)} = \mathbf{M}\left(e^{it\xi_1}e^{it\xi_2}\right) = \\ = \mathbf{M}e^{it\xi_1} \cdot \mathbf{M}e^{it\xi_2} = \phi_1(t) \cdot \phi_2(t)$$

через незалежність випадкових величин $e^{it\xi_1}$ і $e^{it\xi_2}$

Властивості 5) — 7) залишимо без доведення. Тільки формально обчислимо (для властивості 6))

$$\phi^{(n)}(t) = \int\limits_{-\infty}^{+\infty} (ix)^n e^{itx} f(x)dx.$$

Тому

$$\phi^{(n)}(0) = i^n \int\limits_{-\infty}^{+\infty} x^n f(x)dx = i^n \mathbf{M}\xi^n.$$

Зауваження. Для того, щоб неперервна функція $f(t)$ задана на R була характеристичною функцією деякого розподілу, необхідно і досить, щоб вона задовольняла розглянуті вище властивості 1) і 3) (теорема Бохнера-Хінчина).

2.5.2 Характеристична функція нормального розподілу

Нехай випадкова величина ξ має нормальний розподіл $N(a, \sigma^2)$. Тоді її характеристична функція визначається як

$$\mathbf{M}e^{it\xi} = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} e^{itx} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} dx = \\ = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\left(\frac{(x-a)^2}{\sigma^2} - 2it(x-a)\right) + ita\right\} dx = \\ = e^{ita} \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} \exp\left\{-\frac{(x-a-it\sigma^2)^2}{2\sigma^2} - \frac{t^2\sigma^2}{2}\right\} dx = \\ = e^{ita-\frac{t^2\sigma^2}{2}} \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} \exp\left\{-\frac{(x-a-it\sigma^2)^2}{2\sigma^2}\right\} dx = \\ = e^{ita-\frac{t^2\sigma^2}{2}}.$$

Тобто $\phi(t) = e^{ita-\frac{t^2\sigma^2}{2}}$.

Зауваження. Для наступних застосувань випишемо характеристичну функцію стандартного нормального розподілу: $\phi(t) = e^{-\frac{t^2}{2}}$.

На завершення розглянемо корисне твердження: сума двох незалежних нормально розподілених випадкових величин нормально розподілена. Дійсно, використовуючи одну з властивостей характеристичних функцій матимемо, що характеристична функція суми незалежних випадкових величин з розподілами $N(a_1, \sigma_1^2)$ і $N(a_2, \sigma_2^2)$ така

$$\phi(t) = e^{ita_1-\frac{t^2\sigma_1^2}{2}} \cdot e^{ita_2-\frac{t^2\sigma_2^2}{2}} = \\ = e^{it(a_1+a_2)-\frac{t^2(\sigma_1^2+\sigma_2^2)}{2}},$$

що є характеристичною функцією номального розподілу $N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$. Оскільки характеристична функція однозначно визначає розподіл, то згадане твердження доведено.

Легко також встановити, що добуток нормально розподіленої випадкової величини на число має нормальний розподіл.

Дійсно, його характеристична функція

$$\mathbf{M}e^{it(c\xi)} = \mathbf{M}e^{i(tc)\xi} = e^{i(tc)a-\frac{(tc)^2\sigma^2}{2}} = \\ = e^{it(ac)-\frac{t^2(\sigma c)^2}{2}}$$

є характеристичною функцією розподілу $N(ac, (\sigma c)^2)$.

Звідси можна зробити висновок, що лінійна комбінація незалежних в сукупності нормально розподілених випадкових величин має нормальний розподіл.

Вправи до розділу 2

0.10. Закон розподілу випадкової величини ξ дискретного типу заданий таблицею

x_i	1	2	3	4
p_i	1/16	1/4	1/2	3/16

Побудувати графік функції розподілу. Знайти $\mathbf{M}\xi$, $\mathbf{D}\xi$, $\mathbf{P}(\xi > 2)$, $\mathbf{P}(\xi \geq 3)$, $\mathbf{P}(1 < \xi < 3)$.

0.11. Випадкова величина ξ розподілена за законом, що визначається щільністю розподілу виду

$$f_\xi(x) = \begin{cases} C\ \cos x, & \text{якщо} \quad -\pi/2 \leq x \leq \pi/2, \\ 0, & \text{якщо} \quad |x| > \pi/2. \end{cases}$$

Знайти константу C, обчислити $\mathbf{P}(|\xi| < \pi/4)$, $\mathbf{M}\xi$, $\mathbf{D}\xi$.

0.12. Стріляють в ціль три рази. Влучення при окремих пострілах — незалежні події з ймовірністю $\frac{2}{3}$. Нехай ξ — кількість влучень при трьох пострілах. Знайти розподіл ξ, $\mathbf{M}\xi$, $\mathbf{D}\xi$ (*схема Бернуллі*).

0.13. Гральний кубик підкидають до першої появи шестірки. Нехай ξ — кількість підкидань. Знайти розподіл ξ, $\mathbf{M}\xi$, $\mathbf{D}\xi$, $\mathbf{P}(\xi > 10)$ (*геометричний розподіл*).

0.14. Автобуси йдуть з інтервалом 5 хвилин. Вважаючи, що випадкова величина ξ — час чекання автобуса на зупинці — розподілена рівномірно на вказаному інтервалі, знайти функцію розподілу ξ, середній час чекання, дисперсію часу чекання та ймовірність того, що час чекання перевищить 3 хв.

0.15. Випадкова величина ξ розподілена за законом Коші, що задається функцією розподілу виду

$$F_\xi(x) = b + c\ \mathrm{arctg}\frac{x}{a} \quad \text{при} \ -\infty < x < +\infty\ .$$

Вибрати коефіцієнти a, b і c так, щоб дана функція була функцією розподілу неперервної випадкової величини. Обчислити щільність розподілу Коші. Чи існує математичне сподівання цього розподілу? Знайти моду, медіану і квантиль t_p порядку $p = 0.75$ розподілу Коші.

0.16. На контроль поступила партія деталей з цеху. Відомо, що 5% всіх деталей не задовольняють вимоги стандарту. Скільки потрібно випробувати деталей, щоб з ймовірністю не менше 0.95 виявити хоча б одну нестандартну деталь?

0.17. При випробуванні легованої сталі на вміст вуглецю ймовірність того, що у випадково взятій пробі процент вуглецю перевищить допустимий рівень, дорівнює $p = 0.01$. Скільки в середньому необхідно випробувати зразків, щоб з ймовірністю $p = 0.95$ вказаний ефект спостерігався хоча б один раз.

0.18. Випадкова величина ξ розподілена за законом $N(m, \sigma^2)$. Користуючись таблицею функції нормального розподілу, обчислити ймовірність p_k того, що відхилення величини ξ від її математичного сподівання не перевищить величини $k\sigma$ (відповідь одержати для трьох значень $k = 1, 2, 3$).

0.19. Хімічний завод виготовляє сірчану кислоту номінальної густини 1.84 г/см^3. В результаті статистичних досліджень виявлено, що практично 99.9% всіх реактивів, що випускаються, мають густину в інтервалі (1.82; 1.86). Знайти ймовірність того, що кислота задовольняє вимоги стандарту, якщо для цього досить, щоб її густина не відхилялась від номіналу більше, ніж на 0.01 г/см^3.

0.20. Середнє число викликів, що поступають на АТС за хвилину, дорівнює 120. Знайти ймовірності таких подій: $A =${за дві секунди на АТС не поступить жодного виклику}, $B =${за дві секунди на АТС поступить менше двох викликів}, $C =${за одну секунду на АТС поступить три виклики}, $D =${за три секунди на АТС поступить не менше трьох викликів}.

0.21. Деталь, що виготовлена автоматом, вважається якісною, якщо відхилення ξ деякого розміру від номіналу не перевищує 10 мкм. Точність виготовлення деталей характеризується стандартним відхиленням σ. Вважаючи, що для даної технології $\sigma = 5$ мкм і ξ нормально розподілена, визначити: а) скільки процентів якісних деталей виготовляє автомат; б) якою повинна бути точність виготовлення, щоб процент якісних деталей підвищився до 98.

0.22. Випадкова величина ξ розподілена за показниковим законом з параметром λ. Знайти щільності розподілу випадкових величин $\eta = \sqrt{\xi}$, $\psi = \xi^2$, $\zeta = 1 - e^{-\lambda\xi}$.

0.23. Щільність розподілу випадкового вектора (ξ, η) має вигляд:

$$f_{\xi,\eta}(x, y) = \begin{cases} c(x + y), & \text{якщо} \ \ 0 \leq x \leq 1, \ \ 0 \leq y \leq 1, \\ 0, & \text{в інших випадках.} \end{cases}$$

а) Визначити константу c і обчислити $\mathbf{P}(\xi + \eta < 1)$.

б) Визначити безумовну щільність розподілу компоненти η і встановити, залежні компоненти ξ і η чи ні.

в) Знайти $\mathbf{M}\xi$ та $\mathbf{M}\eta$.

0.24. Задані такі характеристики двовимірного нормального вектора (ξ, η): $\mathbf{M}\xi = -2$, $\mathbf{M}\eta = 3$ і коваріаційна матриця $K = \begin{pmatrix} 16 & 0 \\ 0 & 25 \end{pmatrix}$. Записати вираз для щільності розподілу $f_{\xi,\eta}(x, y)$ та обчислити ймовірність попадання в головний еліпс розсіювання з півосями $a = 2\sigma_\xi$, $b = 2\sigma_\eta$.

Розділ 3

Граничні теореми теорії ймовірностей

3.1 Збіжність послідовностей випадкових величин. Закони великих чисел

3.1.1 Нерівності Чебишова

Граничні теореми пов'язують викладену раніше теорію та її застосування. Одна з таких теорем — теорема Пуассона — вже була розглянута. В цьому розділі ми розглянемо більш загальні теореми. А почнемо з нерівностей, які будуть відігравати важливу роль, як в доведенні граничних теорем так і в розв'язанні багатьох практичних задач.

Зафіксуємо деякий ймовірнісний простір $(\Omega, \mathcal{F}, \mathbf{P})$; будемо вважати, що всі випадкові величини, які будуть розглядатися, задані на цьому просторі.

Теорема 1. (*Узагальнена нерівність Чебишова*). Нехай ξ — довільна випадкова величина і $g(x)$ — невід'ємна парна і неспадна на $[0; +\infty)$ функція. Тоді, якщо існує $\mathbf{M}g(\xi)$, то при всіх $\varepsilon > 0$ таких, що $g(\varepsilon) \neq 0$,

$$\mathbf{P}(|\xi| \geq) \leq \frac{\mathbf{M}g(\xi)}{g(\varepsilon)}.$$

Доведення. Доведемо цю теорему в припущенні, що випадкова величина ξ неперервна. Нехай $f(x)$ — її щільність розподілу. Тоді

$$\mathbf{M}g(\xi) = \int\limits_{-\infty}^{+\infty} g(x)f(x)dx =$$
$$= \int\limits_{-\infty}^{-\varepsilon} g(x)f(x)dx + \int\limits_{-\varepsilon}^{\varepsilon} g(x)f(x)dx + \int\limits_{\varepsilon}^{+\infty} g(x)f(x)dx \geq$$
$$\geq \int\limits_{-\infty}^{-\varepsilon} g(x)f(x)dx + \int\limits_{\varepsilon}^{+\infty} g(x)f(x)dx$$

бо $g(x)f(x) \geq 0$ при всіх $x \in R$. Звідси

$$\mathbf{M}g(\xi) \geq g(\varepsilon) \int\limits_{-\infty}^{-\varepsilon} f(x)dx + g(\varepsilon) \int\limits_{\varepsilon}^{+\infty} f(x)dx,$$

бо функція $g(x)$ парна і $g(x) \geq g(\varepsilon)$ при $x \geq \varepsilon$ (неспадна).

Отже, $\mathbf{M}g(\xi) \geq g(\varepsilon) \cdot \mathbf{P}(|\xi| \geq \varepsilon)$. Звідси і випливає твердження теореми.

Розглянемо наслідки цієї теореми.

Наслідок. (*перша нерівність Чебишова*). Якщо випадкова величина ξ має скінченний момент $\mathbf{M}|\xi|$, то для всіх $\varepsilon > 0$

$$\mathbf{P}(|\xi| \geq \varepsilon) \leq \frac{\mathbf{M}|\xi|}{\varepsilon}.$$

При умові, що $\mathbf{P}(\xi \geq 0) = 1$,

$$\mathbf{P}(\xi \geq \varepsilon) \leq \frac{\mathbf{M}\xi}{\varepsilon}.$$

Доведення. Застосуємо теорему 1 з $g(x) = |x|$.

Наслідок. (*друга нерівність Чебишова в нецентрованій формі*). Якщо існує $\mathbf{M}\xi^2$ (чи $\mathbf{D}\xi$), то при всіх $\varepsilon > 0$

$$\mathbf{P}(|\xi| \geq \varepsilon) \leq \frac{\mathbf{M}\xi^2}{\varepsilon^2}.$$

Доведення. Візьмемо в теоремі 1 функцію $g(x) = x^2$.

Наслідок. (*друга нерівність Чебишова в центрованій формі*). Якщо існує $\mathbf{M}\xi^2$ (чи $\mathbf{D}\xi$), то при всіх $\varepsilon > 0$

$$\mathbf{P}(|\xi - \mathbf{M}\xi| \geq \varepsilon) \leq \frac{\mathbf{D}\xi}{\varepsilon^2}.$$

Доведення. Досить застосувати до випадкової величини $\xi - \mathbf{M}\xi$ і функції $g(x) = x^2$ теорему 1.

Розглянемо приклад застосування нерівності Чебишова. **Приклад 1.** Скільки вимірів потрібно зробити, щоб середнє арифметичне результатів дало вимірювану величину з точністю 0.05 та надійністю 90%, якщо дисперсії (задаються точністю приладу) результатів вимірів (як випадкових величин) не більші ніж 0.2?

Перш ніж розв'язувати цю задачу, дамо деякі пояснення. Припустимо, що вимірюється деяка величина, значення якої l. При кожному вимірі неминучі помилки. Але ми будемо вважати, що всі вони випадкові, тобто, систематичні помилки відсутні. Ці помилки призводять до того, що в результаті вимірів ми одержимо значення n випадкових величин $\xi_1, \xi_2, \ldots, \xi_n$. Природно вважати їх однаково розподіленими з $\mathbf{M}\xi_i = l$ $(i = \overline{1,n})$ та незалежними. Тому $\frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{M}\xi_i = l$.

Якщо $\mathbf{P}\left(\left|\frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{M}\xi_i - l\right| < \varepsilon\right) \geq \gamma$, то кажуть, що величина (випадкова) $\frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{M}\xi_i$ наближає значення l з точністю ε та надійністю γ (іноді надійність виражають у процентах).

В нашому випадку $\varepsilon = 0.05$, $\gamma = 0.9$. Скористаємося другою нерівністю Чебишова, в центрованій формі, для випадкової величини $\eta = \frac{1}{n}\sum\limits_{i=1}^{n}\xi_i$.

$$\mathbf{M}\eta = \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i = l, \quad \mathbf{D}\eta = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{D}\xi_i \leq \frac{0.2 \cdot n}{n^2} = \frac{0.2}{n}.$$

Тому

$$\mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i - l\right| < \varepsilon\right) =$$
$$= 1 - \mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i - l\right| \geq \varepsilon\right) \geq 1 - \frac{0.2}{n^2 \cdot \varepsilon^2}.$$

Знайдемо n з умови $1 - \frac{0.2}{n^2 \cdot \varepsilon^2} \geq \gamma$. Отже,

$$n \geq \frac{0.2}{(1-\gamma)\varepsilon^2} = \frac{0.2}{0.1 \cdot 0.05^2} = 800.$$

Тому досить зробити 800 вимірів.

3.1.2 Закони великих чисел

Розглянемо послідовність випадкових величин $\{\xi_n\}$, $(n \geq 1)$, що мають скінченні математичні сподівання. Нас цікавитиме відповідь на питання: чи будуть середні арифметичні $\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ і $\frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{M}\xi_i$ в якомусь розумінні наближатися одне до одного при рості n. З таким питанням ми вже зустрічалися в прикладі попереднього пункта. Відповіді на це питання і становлять суть, так званих, *законів великих чисел.*

Теорема 2. (*Чебишова*). Нехай для послідовності незалежних випадкових величин $\{\xi_n\}$, $(n \geq 1)$ існують $\mathbf{D}\xi_n$ і $\mathbf{D}\xi_n \leq C$ при всіх n. Тоді

$$\mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i\right| \geq \varepsilon\right) \to 0,\ n \to \infty$$

Доведення. Розглянемо випадкові величини $\eta_n = \frac{1}{n}\sum\limits_{i=1}^{n}\xi_i$.

$$\mathbf{M}\eta_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i, \quad \mathbf{D}\eta_n = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{D}\xi_i \leq \frac{nC}{n^2} = \frac{C}{n}$$

Тому за другою нерівністю Чебишова в центрованій формі

$$\mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i\right| \geq \varepsilon\right) \leq \frac{\mathbf{D}\eta_n}{\varepsilon^2} = \frac{C}{n\varepsilon^2}$$

Оскільки $\lim\limits_{n\to\infty}\frac{C}{n\varepsilon^2} = 0$, а ймовірність невід'ємна, то має місце твердження теореми.

Зауваження.

1. Твердження теореми Чебишова називається *законом великих чисел.*

2. З доведення теореми Чебишова видно, що якщо відмовитись від незалежності випадкових величин ξ_n, то для виконаня закону великих чисел досить вимагати щоб $\frac{1}{n^2}\mathbf{D}\left(\sum\limits_{i=1}^{n}\xi_i\right) \to 0$, $n \to \infty$ (*теорема Маркова*).

В теоремі Чебишова ми зустрілися з однією з форм збіжності послідовності випадкових величин, а саме із *збіжністю за ймовірністю.*

Означення. Послідовність випадкових величин $\{\xi_n\}$, $(n \geq 1)$ збігається за ймовірністю до випадкової величини ξ, якщо для всіх $\varepsilon > 0$

$$\mathbf{P}(|\xi_n - \xi| \geq \varepsilon) \to 0,\ n \to \infty.$$

Позначається це так: $\xi_n \xrightarrow{\mathbf{P}} \xi$, $n \to \infty$.

Таким чином закон великих чисел стверджує, що різниця $\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i - \frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{M}\xi_i$ збігається за ймовірністю до 0.

Наведемо один наслідок теореми Чебишова.

Наслідок. (*теорема Бернуллі*). Відносна частота успіхів в n випробуваннях за схемою Бернуллі збігається за ймовірністю при $n \to \infty$ до ймовірності успіху в одному випробуванні.

Доведення. Нехай X_k — індикатор успіху в k-ому випробуванні. Тоді $\nu_n = \frac{1}{n}\sum\limits_{k=1}^{n}X_k$ — відносна частота успіху в n випробуваннях. Нехай p — ймовірність успіху в кожному випробуванні. Тоді

$$\mathbf{P}(X_k = 1) = p, \quad \mathbf{P}(X_k = 0) = 1 - p,$$
$$\mathbf{M}X_k = p, \quad \mathbf{D}X_k = p - p^2 = p(1-p).$$

Звідси $\frac{1}{n}\sum\limits_{k=1}^{n}\mathbf{M}X_k = p$ і за теоремою Чебишова

$$\mathbf{P}(|\nu_n - p| \geq \varepsilon) \to 0,\ n \to \infty,$$

або $\nu_n \xrightarrow{\mathbf{P}} p$, $n \to \infty$.

Зауваження. Теорема Бернуллі оправдує, так зване, *емпіричне* чи частотне означення ймовірності. Коли ймовірністю випадкової події вважається число до якого наближається відносна частота появи цієї події при великій кількості незалежних повторень експерименту в незмінних умовах.

Збіжність за ймовірністю $\xi_n \xrightarrow{\mathbf{P}} \xi$, $n \to \infty$, яка виникає, наприклад, в законі великих чисел, ще не забезпечує рівності $\lim\limits_{n\to\infty}\xi_n = \xi$ на всьому Ω, або хоч на його частині з ймовірністю 1. Більш сильним поняттям збіжності є таке.

Означення. Послідовність випадкових величин $\{\xi_n\}$, $(n \geq 1)$ *збігаєтьcя з ймовірністю 1* (майже напевно) до випадкової величини ξ, якщо $\mathbf{P}(\lim\limits_{n\to\infty}\xi_n = \xi) = 1$.

Позначається це так $\xi_n \xrightarrow{\mathbf{P}=1} \xi$, $n \to \infty$, або $\xi_n \xrightarrow{\text{м.н.}} \xi$, $n \to \infty$.

Твердження, аналогічне закону великих чисел, в якому збіжність за ймовірністю замінене на збіжність з ймовірністю 1, називається *підсиленим законом великих чисел.* Сформулюємо відповідну теорему.

Теорема 3. (*Колмогорова*). Нехай $\{\xi_n\}$, $(n \geq 1)$ послідовність незалежних випадкових величин, для яких існує $\mathbf{D}\xi_n$. Якщо $\sum\limits_{k=1}^{\infty}\frac{1}{k^2}\mathbf{D}\xi_k < +\infty$, то

$$\mathbf{P}\left(\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i\right) = 0\right) = 1.$$

Залишимо цю теорему без доведення, яке можна знайти, наприклад, в [1].

Наслідок. (*теорема Бореля*). Відносна частота успіхів в n випробуваннях за схемою Бернуллі збігається з ймовірністю 1 при $n \to \infty$ до ймовірності успіху в одному випробовуванні.

Доведення. Розглянувши ті ж випадккові величини X_k, що і в доведенні теореми Бернуллі, запишемо

$$\sum_{k=1}^{\infty}\frac{1}{k^2}\mathbf{D}X_k = \sum_{k=1}^{\infty}\frac{p(1-p)}{k^2} < +\infty.$$

Тому твердження теореми Бореля випливає з теореми Колмогорова.

Теорема Бореля ще раз свідчить на користь емпіричного поняття ймовірності.

3.2 Центральна гранична теорема

3.2.1 Центральна гранична теорема

Цей параграф присвячений одному з найбільш важливих тверджень теорії ймовірностей: при досить широких припущеннях суми великої кількості незалежних малих випадкових доданків мають розподіл близький до нормального. Цей результат є, наприклад, теоретичною основою застосування нормального розподілу при розв'язуванні багатьох практичних задач. Так помилки вимірів, що дають вимірювальні прилади, швидкості руху молекул газу мають розподіл, що мало відрізняється від нормального.

Розглянемо одну з найпростіших теорем, що узаконює згадане твердження.

Теорема 1. (*центральна гранична теорема*)
Нехай $\{\xi_n\}$, $(n \geq 1)$ — послідовність незалежних однаково розподілених випадкових величин із скінченною дисперсією ($\mathbf{D}\xi_k = \sigma^2 > 0$) і $s_n = \sum\limits_{k=1}^{n} \xi_k$. Тоді при $n \to \infty$ для всіх $x \in R$

$$\mathbf{P}\left(\frac{s_n - \mathbf{M}s_n}{\sqrt{\mathbf{D}s_n}} < x\right) \to \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt. \tag{3.1}$$

Доведення. Нехай $\mathbf{M}\xi_k = a$. Тоді

$$\mathbf{M}s_n = \sum_{k=1}^{n} \mathbf{M}\xi_k = n \cdot a, \quad \mathbf{D}s_n = \sum_{k=1}^{n} \mathbf{D}\xi_k = n \cdot \sigma^2.$$

Розглянемо випадкові величини

$$\psi_n = \frac{s_n - \mathbf{M}s_n}{\sqrt{\mathbf{D}s_n}} = \frac{\sum\limits_{k=1}^{n} \xi_k - n \cdot a}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{\xi_k - a}{\sigma}$$

і $\eta_k = \frac{\xi_k - a}{\sigma}$. Зауважимо, що $\psi_n = \frac{1}{\sqrt{n}} \sum\limits_{k=1}^{n} \eta_k$.

Оскільки

$$\mathbf{M}\eta_k = \tfrac{1}{\sigma}(\mathbf{M}\xi_k - a) = 0,$$
$$\mathbf{D}\eta_k = \tfrac{1}{\sigma^2}\mathbf{D}(\xi_k - a) = \tfrac{1}{\sigma^2}\mathbf{M}(\xi_k - a)^2 = \tfrac{1}{\sigma^2}\mathbf{D}\xi_k = 1,$$

то характеристичні функції випадкових величин η_k

$$\phi(t) = 1 - \frac{t^2}{2} + o(t^2).$$

Тому характеристичні функції випадкових величин ψ_n

$$\phi_{\psi_n}(t) = \left(\phi\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n.$$

Оскільки

$$\lim_{n\to\infty} \phi_{\psi_n}(t) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n = e^{-t^2/2},$$

то при $n \to \infty$ $\mathbf{P}(\psi_n < x) \to \Phi(x)$, бо функція $e^{-t^2/2}$ є характеристичною функцією стандартного нормального розподілу.

Сформулюємо (без доведення) центральну граничну теорему для послідовності неоднаково розподілених випадкових величин.

Теорема 2. (*теорема Ляпунова*). Якщо для послідовності незалежних випадкових величин $\{\xi_n\}$, $(n \geq 1)$, що мають скінченні моменти третього порядку, виконується

$$\lim_{n\to\infty} \frac{\sum\limits_{k=1}^{n} \mathbf{M}|\xi_k - \mathbf{M}\xi_k|^3}{\left(\sum\limits_{k=1}^{n} \mathbf{D}\xi_k\right)^{3/2}} = 0,$$

то має місце співвідношення (3.1).

Зауваження. Зміст умов центральної граничної теореми (в будь-якому варіанті) полягає в тому, що в сумі $\sum\limits_{k=1}^{n} \xi_k$ жоден з доданків не домінує над іншими.

Приклад 1. Розв'яжемо задачу, сформульовану в попередній лекції. Скільки вимірів потрібно зробити, щоб з точністю 0.05 та надійністю 90% середнє арифметичне вимірів дало вимірювану величину, якщо дисперсії результатів вимірів не перевищують 0.2?

Будемо вважати, що результати вимірів ξ_k є незалежними однаково розподіленими випадковими величинами. Нехай $\mathbf{M}\xi_k = a$, $\mathbf{D}\xi_k = \sigma^2$. Тоді при великих n за центральною граничною теоремою

$$\mathbf{P}\left(\frac{\sum\limits_{k=1}^{n} \xi_k - n \cdot a}{\sqrt{n}\sigma} < x\right) \approx \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt.$$

Тому

$$\mathbf{P}\left(\left|\tfrac{1}{n} \sum\limits_{k=1}^{n} \xi_k - a\right| < \varepsilon\right) =$$
$$= \mathbf{P}\left(-\tfrac{\varepsilon\sqrt{n}}{\sigma} < \frac{\frac{1}{n}\sum\limits_{k=1}^{n} \xi_k - a}{\sigma/\sqrt{n}} < \tfrac{\varepsilon\sqrt{n}}{\sigma}\right) \approx$$
$$\approx \Phi\left(\tfrac{\varepsilon\sqrt{n}}{\sigma}\right) - \Phi\left(-\tfrac{\varepsilon\sqrt{n}}{\sigma}\right) = 2\Phi\left(\tfrac{\varepsilon\sqrt{n}}{\sigma}\right) - 1$$

Вибравши n таким, що $2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1 \geq \gamma$, одержимо, що середнє арифметичне результатів вимірів з точністю ε і надійністю γ дає вимірювану величину.

Звідси $\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) \geq \frac{\gamma+1}{2}$ і $\frac{\varepsilon\sqrt{n}}{\sigma} \geq u_{\frac{\gamma+1}{2}}$, де $u_{\frac{\gamma+1}{2}}$ квантиль порядку $\frac{\gamma+1}{2}$ стандартного нормального розподілу. Тому $n \geq \frac{\sigma^2}{\varepsilon^2} u^2_{\frac{\gamma+1}{2}}$. В нашому випадку $\gamma = 0.9$, $\varepsilon = 0.05$, $\sigma^2 = 0.2$ і з таблиці функції стандартного нормального розподілу $u_{0.95} = 1.645$.

Отже, $n \geq \frac{0.2}{0.0025} \cdot (1.645)^2 \approx 220$.

Зауваження. В прикладі ми одержали значення значно менше, ніж у попередній лекції. Це свідчить про те, що центральна гранична теорема є більш точним твердженням, ніж закон великих чисел чи нерівність Чебишова що була застосована раніше.

3.2.2 Теореми Муавра-Лапласа

Розглянемо деякі наслідки центральної граничної теореми стосовно схеми Бернуллі.

Теорема 3. (*інтегральна теорема Муавра-Лапласа*).
Нехай ν_n — кількість успіхів в n випробовуваннях за схемою Бернуллі з ймовірністю успіху в кожному випробовуванні p $(0 < p < 1)$. Тоді при $a < b$

$$\lim_{n\to\infty} \mathbf{P}\left(a \leq \frac{\nu_n - n \cdot p}{\sqrt{np(1-p)}} < b\right) = \Phi(b) - \Phi(a)$$

(тут, як і раніше, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt$).

Доведення. Випадкові величини $\nu_n = \sum\limits_{k=1}^{n} X_k$, де X_k — індикатор успіху в k-ому випробуванні (незалежні однаково розподілені випадкові величини).

$$\mathbf{M}\nu_n = np, \quad \mathbf{D}\nu_n = np(1-p).$$

Тому з центральної граничної теореми випливає твердження даної теореми, бо випадкові величини $\frac{\nu_n - n\cdot p}{\sqrt{np(1-p)}}$ мають граничний стандартний нормальний розподіл.

Зауваження. Можна довести, що

$$\mathbf{P}(m_1 \le \nu_n < m_2) = \Phi\left(\frac{m_2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{np(1-p)}}\right) + \\ + O\left(\frac{1}{\sqrt{np(1-p)}}\right), \ n \to \infty. \quad (3.2)$$

Це твердження також називають інтегральною теоремою Муавра-Лапласа.

Звідси одержимо таку теорему.

Теорема 4. (*локальна теорема Муавра-Лапласа*). Для ν_n — кількості успіхів в n випробовуваннях за схемою Бернуллі

$$\mathbf{P}(\nu_n = m) = \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{(m-np)^2}{2np(1-p)}\right\} + \\ + O\left(\frac{1}{\sqrt{np(1-p)}}\right), \ n \to \infty.$$

Доведення. Оскільки

$$\mathbf{P}(\nu_n = m) = \mathbf{P}(m \le \nu_n < m+1),$$

то з (3.2) та диференційовності функції $\Phi(x)$

$$\mathbf{P}(\nu_n = m) = \Phi\left(\frac{m-np}{\sqrt{np(1-p)}} + \frac{1}{\sqrt{np(1-p)}}\right) - \\ -\Phi\left(\frac{m-np}{\sqrt{np(1-p)}}\right) + O\left(\frac{1}{\sqrt{np(1-p)}}\right) = \\ = \phi\left(\frac{m-np}{\sqrt{np(1-p)}}\right) \cdot \frac{1}{\sqrt{np(1-p)}} + o\left(\frac{1}{\sqrt{np(1-p)}}\right) + \\ +O\left(\frac{1}{\sqrt{np(1-p)}}\right) = \\ = \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{(m-np)^2}{2np(1-p)}\right\} + O\left(\frac{1}{\sqrt{np(1-p)}}\right),$$

бо $\Phi'(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Зауваження. При великих n можна використовувати наближені рівності

$$\mathbf{P}(m_1 \le \nu_n < m_2) \approx \Phi\left(\frac{m_2-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{m_1-np}{\sqrt{np(1-p)}}\right) \quad (3.3)$$

$$\mathbf{P}(\nu_n = m) \approx \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{(m-np)^2}{2np(1-p)}\right\} \quad (3.4)$$

Слід зауважити, що рівність (3.4) може бути застосована і у випадку досить великих p на відміну від того, що в аналогічному випадку відома нам формула Пуассона вимагає малості параметра p. Проте, якщо ймовірність p близька до 0 чи 1, то для підвищення точності слід використовувати теорему Пуассона.

Вправи до розділу 3

0.25. В страховій компанії застраховано 10000 автомобілів. Ймовірність поломки кожного автомобіля в результаті аварії дорівнює 0.006. Кожен власник застрахованого автомобіля платить в рік 12 грн. страхових і в результаті аварії одержує від компанії 1000 грн. Знайти ймовірність подій $A = \{$за рік страхова компанія зазнає збитків$\}$, $B_m = \{$страхова компанія одержить прибуток не менше m грн.$\}$, якщо $m = 40000$, 60000, 80000.

0.26. Випадкова величина ξ — результат виміру деякої фізичної величини, закон розподілу якої невідомий. Визначити, яку максимально можливу відносну точність виміру можна гарантувати з ймовірністю, не меншою 0.95, при таких даних: а) відомо, що $\mathbf{M}\xi = 0.1$, $\sigma(\xi) = 0.02$ і проводиться один вимір; б) проводиться 5 вимірів і ξ є середнім арифметичним одержаних значень; в) проводиться 100 вимірів і ξ є середнім арифметичним одержаних значень та вважається можливим користуватися граничним розподілом за центральною граничною теоремою.

0.27. Для деякого автопарку середнє число автобусів, що відправляються в ремонт після місяця експлуатації на міських лініях, дорівнює 5. Оцінити ймовірність події $= \{$після закінчення місяця буде відправлено в ремонт менше 15 автобусів$\}$. Розглянути два випадки: а) дисперсія не відома, б) дисперсія дорівнює 4.

0.28. З допомогою нерівності Чебишева оцінити ймовірності $p_k = \mathbf{P}\{|\xi - a| \le k\sigma\}$ для k=1, 2, 3, якщо ξ розподілена за законом $N(a, \sigma^2)$.

0.29. Ймовірність народження хлопчика $p = 0.512$. Вважаючи, що можна застосувати локальну та інтегральну теореми Муавра — Лапласа, визначити ймовірності подій: $= \{$серед 100 новонароджених буде 51 хлопчик$\}$, $= \{$серед 100 новонароджених буде більше хлопчиків, ніж дівчаток$\}$, $= \{$різниця між кількістю хлопчиків і дівчаток із 100 новонароджених не перевищує 10$\}$.

0.30. [1] Визначити, з якою надійністю середнє арифметичне вимірів дає вимірювану величину з точністю 0.1, якщо зроблено 500 вимірів і дисперсія випадкових величин (результатів вимірів) дорівнює 0.3.

0.31. Визначити, з якою точністю середнє арифметичне вимірів дає вимірювану величину, якщо вимірів зроблено 400, надійність результата 80% і дисперсії випадкових величин дорівнюють 0.04.

0.32. Скільки вимірів потрібно зробити, щоб середнє арифметичне їх дало вимірювану величину з точністю 0.05 і надійністю 90%, якщо дисперсії випадкових величин дорівнюють 0.2?

0.33. Проводяться послідовні випробування за схемою Бернуллі. Ймовірність появи події A в одному випробуванні $p = 0.6$. Вважаючи можливим застосовувати граничні теореми Муавра — Лапласа, обчислити ймовірності таких подій: $B = \{$подія A відбудеться в більшості з 60 випробувань$\}$, $C = \{$кількість успішних випробувань (відбулась подія A) знаходиться між 30 і 42$\}$, $D = \{$подія A відбудеться в 36 з 60 випробувань$\}$.

0.34. В ящику знаходяться білі і чорні кульки у співвідношенні 3:2. Проводяться послідовні досліди по витягуванню однієї кульки з поверненням, причому кожного разу фіксується колір вибраної кульки. Яким повинна бути мінімальна кількість витягувань, при якій з ймовірністю, не меншою 0.9948, можна чекати, що відхилення відносної частоти появи білої кульки від ймовірності її появи в одному досліді не перевищує величини $\epsilon = 0.05$?

[1] Розв'язати задачі 30, 31, 32 з допомогою нерівності Чебишева та з допомогою центральної граничної теореми. Порівняти результати.

Додаток

Елементи комбінаторики

Очевидним є те, що у випадку скінченної множини Ω міра (ймовірність) кожної події повинна бути тісно пов'язана з кількістю елементарних подій, що сприяють даній події. Розглянемо тому деякі найпростіші поняття *комбінаторики* — галузі математики,що займається вивченням скінченних множин. В межах цього пункту будемо переважно говорити про множини без зв'язку їх із подіями.

Для початку розглянемо два простих правила, які називаються *правилами суми і добутку.*

Якщо деякий об'єкт α можна вибрати m різними способами, а об'єкт β — n різними способами, причому ніякий вибір α не збігається з жодним з виборів β, то один з об'єктів α або β можна вибрати $m + n$ способами (правило суми).

Доведення. Позначивши через $|A|$ кількість елементів множини A, правило суми можемо записати у вигляді: якщо $|A| = m$, $|B| = n$ і $A \cap B = \emptyset$, то $|A \cup B| = m + n$, що очевидно.

Якщо деякий об'єкт α можна вибрати m різними способами, і при кожному виборі об'єкта α об'єкт β можна вибрати n різними способами, то вибір впорядкованої пари об'єктів $(\alpha,\, \beta)$ можна здійснити $m \cdot n$ способами (правило добутку — *основний принцип комбінаторики*).

Доведення. Позначимо різні способи вибору об'єкта α через $a_1, \ldots, a_m$, а різні способи вибору об'єкта β при виборі α способом a_i через $b_{i1}, \ldots, b_{in}$. Тоді всі можливі способи вибрати впорядковану пару $(\alpha,\, \beta)$ можна скласти в таблицю

$$\begin{array}{cccc}
(a_1,\, b_{11}) & (a_1,\, b_{12}) & \ldots & (a_1,\, b_{1n}) \\
(a_2,\, b_{21}) & (a_2,\, b_{22}) & \ldots & (a_2,\, b_{2n}) \\
\ldots & \ldots & \ldots & \ldots \\
(a_m,\, b_{m1}) & (a_m,\, b_{m2}) & \ldots & (a_m,\, b_{mn})
\end{array}$$

Така таблиця має $m \cdot n$ елементів.

Означення. Розміщенням з n по k ($k \le n$) називають будь-яку впорядковану k-елементну підмножину n-елементної множини.

Зауваження. n-елементна множина називається впорядкованою, якщо кожному її елементу покладено у відповідність одне з чисел 1, 2,..., n, що є його порядковим номером, причому, різні елементи занумеровані різними числами.

Очевидно, що розміщення є різними, якщо або вони складаються з різних елементів, або відрізняються їх порядком.

Кількість розміщень з n по k будемо позначати A_n^k.

Теорема 1. Для будь-яких цілих чисел n і k ($0 \le k \le n$)

$$A_n^k = \frac{n!}{(n-k)!}. \tag{D.1}$$

Доведення. Будь-яке розміщення із n по k можна одержати так: вибрати один елемент із n на перше місце в розміщенні (n способів), потім вибрати один елемент із $n-1$ на друге місце ($n-1$ спосіб) і т. д., і, нарешті, вибрати один елемент із $n-k+1$ на k-те місце ($n-k+1$ спосіб). Зо основним принципом комбінаторики побудувати розміщення із n по k можна

$$A_n^k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

різними способами.

Зауваження. Нагадаємо, що за означенням

$$n! = \begin{cases} 1, & \text{якщо } n = 0 \\ 1 \cdot 2 \cdot \ldots \cdot n, & \text{якщо } n \in N \end{cases}$$

Означення. Перестановкою з n елементів називається будь-яка впорядкована n-елементна множина.

Зауваження. Очевидно, що перестановка з n елементів є розміщенням із n по n.

Кількість перестановок із n елементів позначається через P_n. Оскільки $P_n = A_n^n$, то має місце така теорема.

Теорема 2. Кількість перестановок із n елементів дорівнює

$$P_n = n!. \tag{D.2}$$

Означення. Комбінацією з n елементів по k $(0 \le k \le n)$ називається будь-яка k-елементна підмножина n-елементної множини.

Зауваження. Комбінації, на відміну від розміщень,— це невпорядковані підмножини даної множини. Тому різні комбінації відрізняються тільки складом елементів, що їх утворюють.

Кількість всіх комбінацій із n елементів по k позначається C_n^k.

Теорема 3. Для будь-яких цілих чисел n і k $(0 \le k \le n)$

$$C_n^k = \frac{n!}{k! \cdot (n-k)!}. \tag{D.3}$$

Доведення. Якщо задано деяку комбінацію із n елементів по k, то впорядковуючи її різними способами, одержимо P_k різних розміщень із n елементів по k. За основним принципом комбінаторики $C_n^k \cdot P_k = A_n^k$.

Тому $C_n^k = \frac{A_n^k}{P_k}$ і з (D.1) та (D.2) матимемо

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

Розглянемо деякі важливі властивості кількості комбінацій.

Теорема 4. Для будь-яких цілих чисел n і k $(0 \le k \le n)$ має місце рівність $C_n^k = C_n^{n-k}$.

Доведення. Твердження теореми очевидно випливає з (D.3).

Теорема 5. Для будь-яких цілих чисел n і k $(0 \le k \le n)$ має місце рівність

$$C_{n+1}^k = C_n^k + C_n^{k-1}. \tag{D.4}$$

Доведення. Використовуючи (D.3) одержимо

$$\begin{gathered} C_n^k + C_n^{k-1} = \tfrac{n!}{k!\cdot(n-k)!} + \tfrac{n!}{(k-1)!\cdot(n-k+1)!} = \\ = \tfrac{n!\cdot(n-k+1)+n!\cdot k}{k!\cdot(n-k+1)!} = \tfrac{n!\cdot(n+1)}{k!\cdot(n-k+1)!} = \\ = \tfrac{(n+1)!}{k!\cdot(n+1-k)!} = C_{n+1}^k. \end{gathered}$$

$$
\begin{array}{ccccccccccccc}
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
. & . & . & . & . & . & . & . & . & . & . & . & .
\end{array}
$$

Табл. D.1:

З формули (D.4) випливає легкий спосіб обчислення чисел C_n^k (*біноміальні коефіцієнти*).

Складемо таблицю (*трикутник Паскаля*) за такими правилами:

- перший рядок складається з двох одиниць;
- кожен наступний рядок починається і закінчується одиницею, кожне з інших чисел записується між двома сусідніми числами попереднього рядка і дорівнює їх сумі (див. таблицю D.1).

В n-ому рядку такої таблиці будуть записані числа C_n^k впорядковані за зростанням k.

Теорема 6. Кількість всіх підмножин n-елементної множини дорівнює $\sum\limits_{k=0}^{n} C_n^k = 2^n$.

Доведення. Нехай r_n — кількість всіх підмножин n-елементної множини. Розглянемо $(n+1)$-елементну множину і зафіксуємо один з її елементів a. Розіб'ємо всі підмножини цієї множини на ті, що містять a і ті, які a не містять. Підмножини, що a не містять є підмножинами n-елементної множини. Тому їх є r_n. Підмножини, що містять a можуть бути одержані з попередніх приєднанням до них елемента a. Тому їх також є r_n. Отже, $r_{n+1} = 2 \cdot r_n$. Оскільки $r_1 = 2$, бо одноелементна множина має дві підмножини (саму себе і $\emptyset$), то $r_n = 2^n$.

Бібліоґрафія

[1] Гихман И.И., Скороход А.В., Ядренко М.Й. Теория вероятностей и математическая статистика. – К.: Вища школа, 1988.

[2] Скороход А.В. Елементи теорій ймовірностей та випадкових процесів. К., Вища школа, 1975.

[3] Скороход А.В. Лекції з теорії випадкових процесів: Навч. посібник. – К.: Либідь, 1990.

[4] Шефтель З.Г. Теорія ймовірностей. К., Вища школа, 1995.

[5] Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. Справочник по теории вероятностей и математической статистике. – М.: Наука, 1985.

[6] Ивашев-Мусатов О.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979.

[7] Сборник задач по математике для втузов. Ч.3. Теория вероятностей и математическая статистика./Под ред. А.В. Ефимова. – М.: Наука, 1990.

[8] Вишневский Л.Д., Гусак Д.В., Погребецкая Т.А., Тер-Саакянц Г.Л. Математическая статистика и случайные процессы: Практикум: Учебное пособие. – К.: Вища школа, 1992.

[9] Гураль І.М., Осипчук М.М. Теорія ймовірностей, випадкові процеси та математична статистика: Завдання для практичних занять. Івано-Франківськ, ІФДТУНГ, 1998.

[10] Методичні вказівки і контрольні завдання з курсу "Вища математика"... За ред. В.І. Горгули. Івано-Франківськ, 1996.

Зміст