

Дискретна ймовірність в курсі дискретної математики

Курінний Г.Ч. Невмержицька О.М. Шугайло О.М.

Травень, 2015

Зміст

1	Основні поняття дискретної ймовірності	3
1.1	Простір подій. Ймовірність події	3
2	Розподіли	14
2.1	Випробування Бернуллі. Розподіл Бернуллі	14
2.2	Рівномірний розподіл	14
2.3	Біномний розподіл, наближена формула Пуасона	19
2.4	Розподіл Пуасона	20
2.5	Гіпергеометричний розподіл	24
2.6	Геометричний розподіл	27
2.7	Від’ємний біномний розподіл	28
2.8	Поліномний розподіл	32
2.9	Статистики Максвела — Больцмана, Бозе — Енштейна, Фермі — Дирака	34
2.10	Локальна та інтегральна теореми Лапласа	35
3	Умовна ймовірність	39
3.1	Незалежні події	39
3.2	Ймовірність послідовності незалежних випробувань	42
3.3	Означення умовної ймовірності.	44

3.4	Формула повної ймовірності	46
3.5	Формули Баєса	50
4	Випадкова величина, ймовірність випадкової величини	56
4.1	Випадкова величина	56
4.2	Ймовірність випадкової величини.	57
4.3	Інтегральна (кумулятивна) функція розподілу ймовірності	60
4.4	Випадок двох величин на одному ймовірнісному просторі	61
5	Характеристики розподілу ймовірності випадкової величини	64
5.1	Нескінченні суми дійсних чисел	64
5.2	Математичне очікування	66
5.3	Розсіювання (дисперсія)	70
6	Загальна алгебра і дискретна ймовірність	73
6.1	Кільце формальних степеневих рядів	73
6.1.1	Означення кільця	73
6.1.2	Дії з рядами та многочленами	76
6.1.3	Породжуюча функція розподілу — генератриса	82
6.2	Напівгрупи	86
6.2.1	Вільна напівгрупа над заданим алфавітом	86
6.2.2	Напівгрупове кільце	87
6.2.3	Приклади знаходження генератриси	89
6.2.4	Напівгрупа стохастичних матриць і ланцюги Маркова	92
6.3	Рекурентні послідовності.	98
6.3.1	Процеси і послідовності	98

6.3.2	Рекурентні співвідношення та послідовності ¹	100
6.3.3	Задача про розорення гравця . . .	102

1 Основні поняття дискретної ймовірності

1.1 Простір подій. Ймовірність події

Теорія ймовірностей має давню історію, тому має свою усталену термінологію, яка склалася до введення новітніх основ математики, задовго до введення нової термінології. Викладання теорії ймовірностей ведеться у тій термінології, яка склалася історично.

Теорія ймовірностей починається із введення так званої генеральної сукупності або простору подійору — це два синонімічні поняття. Відповідником терміну “простір подій” в інших розділах математика є термін “множина”. Простір подій звичайно позначається Ω .

Простір подій складається із елементарних подій. Елементарні події називають також точками ймовірнісного простору. В інших розділах математики кажуть про елемент множини. Елементарні події звичайно позначаються ω , і, відповідно, пишеться $\omega \in \Omega$. Замість слова “підмножина” в теорії ймовірностей використовується термін “подія”.

Первісним поштовхом для розвитку теорії ймовірностей, як вважається, були азартні ігри, в яких потрібно оцінити ступінь упевненості у виграші. В прикладах використовуємо чотири гри — “гра в карти”, “підкидання кістки”, “підкидання монети” та доміно.

Гра в карти здійснюється за допомогою “колоди карт”, яка складається із 52 карт — будемо брати так зва-

¹Лінійні рекурентні співвідношення ретельно розглядаються при вивченні визначників матриць і в комбінаториці. Тому наразі текст має характер нагадування

ну бріджеву колоду. Кожна карта має свою “масть” і “зріст”. Мастей 4 — вони мають позначення $\spadesuit$, $\clubsuit$, $\diamondsuit$, $\heartsuit$, і мають назви

$\spadesuit$ — піка або вино;

$\clubsuit$ — трефа (від французького *le trèfle*, що означає конюшину, трилисник) або жир (в карпатському регіоні так називають букове насіння);

$\diamondsuit$ — бубна;

$\heartsuit$ — чирва або черва.

Зріст карт позначається 2,3,4,5,6,7,8,9,10,В,Д,К,Т, і називається відповідно двійка, трійка, четвірка, п'ятірка, шістка, сімка, вісімка, дев'ятка, десятка, валет, дама, король і туз. Таким чином $\diamondsuit$ Д означає бубнову даму.

Вино та жир — це чорні масті, а бубна та чирва — червоні масті.

Приклад 1.1 *Колода карт — це генеральна сукупність, Вибір однієї карти із колоди — це елементарна подія. Вибір карти певної масті, чи певного зросту — це подія.*

Кістка для підкидання — це куб, на гранях якого зображені числа 1, 2, 3, 4, 5 та 6.

Приклад 1.2 *При падінні кубика одна грань із зображенням певного числа виявляється верхньою (випала певна кількість очок), це і є елементарна подія. Випадіння парної кількості очок — це подія. Генеральна сукупність — це 1,2,3,4,5,6.*

Монета має дві сторони — на одній зображена вартість (цифра), а на другій — герб.

Приклад 1.3 При підкидання монети падає однією стороною вгору це і є елементарна подія. Отже елементарних подій в цьому просторі подій дві — Ц і Г.

При грі в доміно використовують комплект, що складається із 28 кісток, кожна з якої помічена двома числами із множини 0,1,2,3,4,5,6. Кістка може бути помічена двома однаковими числами —такі кістки називають дублями. Кістки симетричні, тобто порядок чисел, що зображені на кістці, не має значення — кістка з позначенням (5,6) і кістка з позначенням (6,5) це одна і та ж кістка. При грі в доміно кістки приставляються одна до другої. Приставити одну кістку до другої можна, коли в них є спільне число в позначенні. Так до дубля (0,0) можна приставити кістку (0,5) а приставити (1,3) неможливо.

Приклад 1.4 Комплект для гри в доміно це генеральна сукупність. Вибір однієї кістки із комплекту — це елементарна подія. Вибір дубля — це подія.

В дискретній теорії простір подій або скінченний, або злічений, тобто елементарних подій або скінченна кількість, або їх можна розташувати у послідовність. Отже можна вважати², що

$$\Omega = \{\omega_0, \omega_1, \omega_2, \dots, \omega_n, \dots\} \quad (1)$$

і ця послідовність може бути скінченною. Порожнім простір подій бути не може, це впливає із обмеження (3)— сума порожньої множини доданків дорівнює нулю.

Буває зручно індексувати елементи дискретного простору подій (елементарні події) цілими числами. І тоді

²При нумерації елементів послідовності в математичному мовному оточенні звичайно перший елемент одержує номер 0. В програмуванні перша комірка пам'яті має адресу 0. Такий підхід дає певні зручності. В позаматематичному оточенні, звичайно, перший елемент (наприклад перший місце в ряду в театрі) має номер 1.

можна можна записати

$$\Omega = \{\dots, \omega_{-2}, \omega_{-1}, \omega_0, \omega_1, \omega_2, \dots, \omega_n, \dots\}.$$

Кожній елементарній події $\omega \in \Omega$ ставиться у відповідність дійсне невід'ємне число

$$0 \leq pr(\omega) \leq 1, \quad (2)$$

яке називають ймовірністю елементарної події ω . Ймовірності елементарних подій підкоряються обмеженню:

$$\sum_{\omega \in \Omega} Pr(\omega) = 1. \quad (3)$$

Ймовірність $Pr(\Omega_1)$ події $\Omega_1 \subseteq \Omega$ визначається правилом

$$Pr(\Omega_1) = \sum_{\omega \in \Omega_1} Pr(\omega).$$

Оскільки простір подій і подія може складатися із нескінченної кількості елементарних подій, то уточнимо, як розуміти суму нескінченної кількості доданків.

У випадку, коли Ω_1 складається із нескінченної кількості елементарних подій, сума $\sum_{\omega \in \Omega_1} Pr(\omega)$ розуміється як точна верхня межа (супремум) скінченних сум: беруться скінченні підмножини $\Omega_2 \subseteq \Omega_1$, шукаються скінченні суми $\sum_{\omega \in \Omega_2} Pr(\omega)$, і верхня межа цих скінченних сум за визначенням дорівнює $\sum_{\omega \in \Omega_1} Pr(\omega)$.

Сума елементів нескінченної послідовності

$$Pr(\omega_0), Pr(\omega_1), Pr(\omega_2), \dots, Pr(\omega_n), \dots$$

розуміється як

$$\sum_{i=0}^{\infty} Pr(\omega_i) = \lim_{n \rightarrow \infty} \sum_{i=0}^n Pr(\omega_i).$$

Користуємося як відомим з курсу математичного аналізу тим, що коли елементи нескінченної послідовності

$$x_0, x_1, \dots, x_n, \dots \quad (4)$$

є дійсними невід’ємними числами, то три умови

- існує сума a елементів послідовності (4);
- існує сума b елементів послідовності (4), що записані в довільному порядку;
- існує точна верхня межа скінченних сум елементів послідовності (4)

рівносильні, тобто коли існує одне із чисел a, b, c , то існують і решта два, причому $a = b = c$.

Визначення 1.1 Функцію

$$Pr : \Omega \rightarrow \mathbb{R}, \quad (5)$$

яка ставить у відповідність кожній елементарній події $\omega \in \Omega$ ймовірність $Pr(\omega)$, що підкоряється обмеженням (2), (3), називається функцією розподілу ймовірності³ у просторі подій. Кажуть також, що функція pr задає розподіл ймовірності, або є розподілом ймовірності. Значення функції розподілу від заданої елементарної події називають ймовірністю елементарної події.

Протстір подій разом із функцією розподілу ймовірності на цьому просторі називають ймовірнісним простором⁴.

Порожньою подія може бути — це неможлива подія, її ймовірність дорівнює 0.

³Таке означення функції розподілу ймовірності використовується в дискретній теорії ймовірностей. В загальних сучасних розділах теорії ймовірності використовують інше означення — див стор. ??

⁴В загальній теорії ймовірності використовують інші підходи — інше означення ймовірнісного простору.

Поряд з позначенням $\Pr(\omega)$ для ймовірності елементарної події і $\Pr(\Omega_1)$ для події з використанням змінної x для елементарної події “змінна x дорівнює ω ” та для події “змінна x належить Ω ” використовуються позначення

$$\Pr(\omega) = \Pr(x = \omega), \quad \Pr(\Omega_1) = \Pr(x \in \Omega_1).$$

Коли шукають ймовірність певної події $\Omega_1 \subseteq \Omega$, то звичайно всі елементарні події із Ω_1 називають сприятливими.

Оскільки події, це підмножини генеральної сукупності, то можна знаходити їх теоретико-множинний перетин і об’єднання. Їх називають, відповідно, добутком і сумою подій. Отже,

коли $A, B \subseteq \Omega$ — дві події в генеральній сукупності Ω , то добуток AB складається із тих елементарних подій, які в один і той же час належать і A і B , а сума $A + B$ складається із тих елементарних подій, які належать хоч одній із подій A, B .

Якщо A — подія, то подія $B = \Omega \setminus A$ називається протилежною подією і позначається $\bar{A}$.

Згідно з введеними позначеннями, $A \setminus B$ для подій $A, B \subseteq \Omega$ буде позначатися $A\bar{B}$, а симетрична різниця подій A, B позначається $A\bar{B} + B\bar{A}$.

Замість того, щоб сказати, що дві події мають порожній перетин, кажуть, що ці дві події не-сумісні. А замість того, щоб сказати, що об’єднання подій є всією генеральною сукупністю, кажуть що ці події утворюють повну групу.

Несумісність подій, які утворюють повну групу, означенням не вимагається, але часто розглядаються саме повні групи несумісних подій.

Те, що в комбінаториці називають сполуками, сполуками з повтореннями, розміщеннями та розміщеннями з повтореннями, в теорії ймовірностей називають вибірками, вибірками з поверненням, впорядкованими вибірками та впорядкованими вибірками з поверненнями.

Теорема 1.1 (Теорема додавання ймовірностей несумісних подій)

Нехай є дві несумісні події $A, B \subseteq \Omega$ на ймовірнісному просторі Ω . Тоді

$$Pr(A + B) = Pr(A) + Pr(B).$$

Доведення. Оскільки $A \cap B = \emptyset$, то

$$Pr(A+B) = \sum_{\omega \in A \cup B} Pr(\omega) = \sum_{\omega \in A} Pr(\omega) + \sum_{\omega \in B} Pr(\omega) = Pr(A) + Pr(B).$$

■

Наслідок 1.1 *Якщо є $n > 2$ взаємно несумісних подій $A_1, A_2, \dots, A_n$, тобто*

$$A_i \subseteq \Omega, \quad A_i \cap A_j = \emptyset, \quad i \neq j, \quad i, j \in \{1, 2, 3, \dots, n\},$$

то

$$pr\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n Pr(A_i).$$

Наслідок обґрунтовується теоремою додавання ймовірностей несумісних подій з використанням методу математичної індукції.

Теорема 1.2 (Теорема додавання ймовірностей сумісних подій)

Для будь-яких двох (можливо, сумісних) подій $A, B \subseteq \Omega$ на ймовірнісному просторі Ω

$$Pr(A + B) = Pr(A) + Pr(B) - Pr(AB). \quad (6)$$

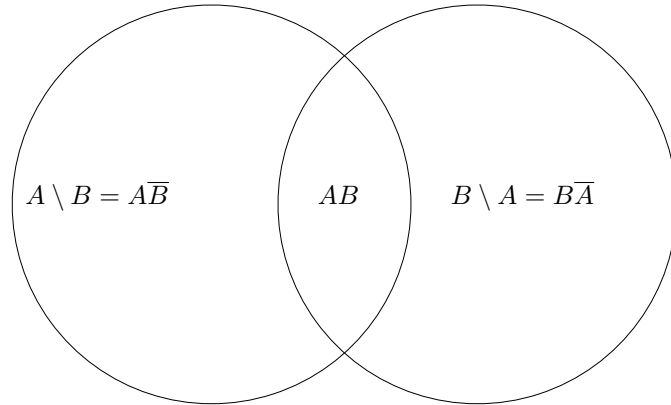


Рис. 1: Суму $C = A + B$ двох сумісних подій A та B можна розглядати як суму трьох несумісних подій $C = A\overline{B} + AB + B\overline{A}$

Доведення. Запишемо подію $C = A + B$ у вигляді суми несумісних подій (див. рис. 1.1)

$$A = A\overline{B} + AB, \quad B = B\overline{A} + AB, \quad C = A\overline{B} + B\overline{A} + AB.$$

Теорема додавання ймовірностей дозволяє записати рівності:

$$\Pr(A) = \Pr(A\overline{B}) + \Pr(AB), \quad \Pr(B) = \Pr(B\overline{A}) + \Pr(AB),$$

$$\Pr(C) = \Pr(A\overline{B}) + \Pr(B\overline{A}) + \Pr(AB).$$

Із виписаних рівностей випливає, що

$$\Pr(C) = \Pr(A\overline{B}) + \Pr(AB) + \Pr(B\overline{A}) + \Pr(AB) - \Pr(AB) = \Pr(A) + \Pr(B) - \Pr(AB)$$

■

Приклад 1.5 Припустимо що є деталі певних кольорів і певних форм, і ймовірність вибрати червону деталь дорівнює 0,3, ймовірність вибрати квадратну деталь дорівнює 0,5, а ймовірність того, що вибрана деталь буде червоною або квадратною, дорівнює 0,6.

Яка ймовірність того, що вибрана деталь виявиться червоною і квадратною

Щоб знайти відповідь на поставлене питання, позначимо через A подію — вибрана деталь є червоною, через B — вибрана деталь є квадратною, і скористаємося формулою (6):

$$\Pr(A + B) = 0,6 = 0,3 + 0,5 - \Pr(AB).$$

Звідси одержуємо відповідь

$$\Pr(AB) = 0,2.$$

Якщо ϵ множина M , то можна розглядати

- 2^M — множину усіх підмножин множини M ;
- M^n — множину послідовностей елементів із M .

Отже, елементарна подія в одному просторі подій може бути подією, або бути послідовністю елементарних подій в іншому просторі подій.

Приклад 1.6 *Нехай є 5 деталей, із яких 2 браковані. Вибір однієї деталі це елементарна подія. Надамо кожній із них ймовірність $\frac{1}{5}$. Вибір бракованої деталі, це подія, не елементарна. Ймовірність цієї події $\frac{2}{5}$. Подія “вибрали не браковану деталь” має ймовірність $\frac{3}{5}$. Тепер ми можемо розглядати простір елементарних подій, який складається із двох елементарних подій: a “вибрали не браковану деталь” і b “вибрали браковану деталь”. Розподіл ймовірностей на цьому просторі елементарних подій буде $\Pr(a) = \frac{3}{5}$, $\Pr(b) = \frac{2}{5}$.*

Приклад 1.7 *Один простір подій складається із можливих видів транспорту, якими рухався студент — електричкою чи автобусом. Другий простір може складатися із можливих оцінок, які одержує студент*

на іспиті — незадовільно, задовільно, добре та відмінно. Можна створити новий простір — декартовий добуток двох створених просторів, в якому елементарними подіями є пара “вид транспорту” і “одержана оцінка”.

Є три основні способи одержати ймовірність елементарної події

Перший спосіб. Елементарна подія ω розглядається як можливий результат певного досліду. Можливі результати досліду — це всі елементи простору подій Ω . За одних і тих же обставин дослід повторюється n разів, де n є великим (в інтуїтивному розумінні) числом. Кількість появи результату ω поділена на n називається частотою появи події ω . Якщо результат ω трапився k раз (обов’язково $0 \leq k \leq n$), то частота появи цього результату в серії дослідів, що складається із n дослідів, дорівнює $\frac{k}{n}$. Далі приписуємо події ω ймовірність, що дорівнює частоті: $\Pr(\omega) = \frac{k}{n}$. Наприклад, підкидаємо кубик 100 разів і одержуємо результат

1 випало 15 разів,
2 випало 20 разів,
3 випало 10 разів,
4 випало 20 разів,
5 випало 20 разів,
6 випало 5 разів.

На основі цієї таблиці приписуємо елементарним подіям ймовірності, що дорівнюють відповідним частотам:

$$\Pr(1) = 0.15, \quad \Pr(2) = 0.2, \quad \Pr(3) = 0.1, \quad \Pr(4) = 0.2, \quad \Pr(5) = 0.2, \quad \Pr(6) = 0.1$$

Ми можемо провести повторну серію дослідів і одержати інші ймовірності. Але є випадки, коли повторну

серію дослідів зробити неможливо. Так коли потрібно знати ймовірність того, що українець доживе до 70 років, то ми візьмемо кількість n померлих останнього року, кількість k померлих цього ж року у віці більше 70 років, і одержимо шукану ймовірність

$$\frac{k}{n}.$$

Звичайно, вибравши інший період часу, ми одержимо іншу ймовірність.

Другий спосіб полягає в тому, що ми в позаматематичний спосіб уявляємо собі багаторазовий дослід, який може закінчитися $\omega_0, \omega_1, \dots, \omega_{n-1}$ результатами, вважаємо ці результати елементарними подіями і приписуємо їм ту границю, до якої прямують частоти, з якими з'являються відповідні результати. Таким чином, ми можемо вважати кубик правильним, і границею, до якої прямують частоти, всі дорівнюють $\frac{1}{6}$.

Стосовно другого способу потрібно мати на увазі, що наша уява про наявність границі, до якої прямують частоти, може бути хибною. Так ми можемо уявити довжину узбережжя Англії як границю, до якої прямує її вимірювання із заданою точністю, коли ця точність прямує до нуля. Як показали дослідження, ця границя не існує, — виміряна довжина прямує до нескінченності, коли точність вимірювання прямує до нуля.

Питання об'єктивної, незалежної ні від часу ні від дослідника, ймовірності елементарної події віднесемо до філософії.

Третій спосіб використовується, коли елементарна подія є просто подією, але в іншому просторі. Тоді її ймовірність можна підрахувати в цьому іншому ймовірнісному просторі. Для прикладу, розглядаємо елементарні події — випадання рівно n , $0 \leq n \leq 10$ гербів

в серії із 10 підкидань монети. Випадання 7 гербів є елементарною подією. Ця елементарна подія є просто подією у просторі, де елементарними подіями ω є серії 10-разових підкидань монети з рівномірним розподілом — $\Pr(\omega) = 2^{-10}$. Тут подія “випало 7 гербів” має ймовірність

$$\Pr(7) = \frac{C_{10}^7}{2^{10}}.$$

2 Розподіли

2.1 Випробування Бернуллі. Розподіл Бернуллі

Розглядається випадок, він називається випробування Бернуллі, коли елементарних подій всього дві — їх можна назвати Успіх та Невдача і відповідно позначити Y та H , з розподілом ймовірностей

$$\Pr(Y) = p, \quad \Pr(H) = q \quad (7)$$

Розподіл ймовірностей (7) на двоелементному просторі подій $\Omega = \{Y, H\}$ називають розподілом Бернуллі.

Приклад 2.1 *Прикладами двоелементного простору подій з відповідним розподілом Бернуллі є випадання герба або цифри при підкиданні монети, попадання пострілу в мішень чи промах, одержання прибутку чи не одержання та подібні.*

2.2 Рівномірний розподіл

Якщо всі елементарні події мають одну і ту ж ймовірність, то такий розподіл називається рівномірним, а елементарні події називаються рівноймовірними.

Якщо генеральна сукупність має n елементів і розподіл рівномірний, то кожна елементарна подія (точка) має ймовірність $\frac{1}{n}$, а подія, що складається із m точок, має ймовірність $\frac{m}{n}$.

Приклад 2.2 *Підкидають три гральні кістки — скажемо, жовту, зелену та червону. Яка ймовірність того, що в сумі випаде 13 очок.*

Оскільки тут мається на увазі рівномірний розподіл, то потрібно знайти кількість n всіх елементарних подій, кількість сприятливих подій m , і розділити кількість сприятливих на кількість усіх елементарних подій.

Ймовірнісний простір містить 6^3 елементарних подій, оскільки підкидання кожної кістки може закінчитися однією із 6 можливостей.

Випишемо ті числа $x \leq y \leq z$, $x + y + z = 13$, що можуть випасти на кістках, і кількість k можливостей розташувати ці три числа на 3 верхніх гранях.

x	y	z	k
1	6	6	3
2	5	6	6
3	4	6	6
3	5	5	3
4	4	5	3

Наведена таблиця показує, що $m = 3 + 6 + 6 + 3 + 3 = 21$. Отже шукана ймовірність дорівнює

$$\frac{m}{n} = \frac{21}{6^3} = \frac{7}{72}.$$

Приклад 2.3 *4 хлопці по черзі миють посуд і у них час від часу розбиваються тарілки. Вважаємо, що ймовірність того, що тарілка буде розбитою хлопцем, одна і та ж для кожного хлопця. Яка ймовірність*

того, що із 4-х розбитих тарілок 3 розіб'є четвертий хопець.

В цьому прикладі потрібно розібратися, що є елементарною подією, і яка ймовірність цієї елементарної події.

Починаємо з того, що розглядаємо елементарні події — тарілка розбивається одним із 4-х хлопців. Цей простір подій Ω_1 4-елементний. Отже кожна з подій $\omega_i \in \Omega_1$, $i = 1, 2, 3, 4$ — тарілка розбита i -м хлопцем, має ймовірність $\frac{1}{4}$;

$$\Pr(\omega_1) = \Pr(\omega_2) = \Pr(\omega_3) = \Pr(\omega_4) = \frac{1}{4}.$$

Тепер розглядаємо нові елементарні події — послідовності довжини 4 елементарних подій із Ω_1 — $(\omega_p, \omega_q, \omega_r, \omega_s)$, де $p, q, r, s, t \in \{1, 2, 3, 4\}$. Новий простір подій Ω_2 має $n = 4^4 = 256$ послідовностей — елементарних подій. Сприятливі події ті, що на трьох місцях мають ω_4 , а четверте місце зайняте $\omega_i, i \in \{1, 2, 3\}$. Таких послідовностей $m = 12$ — їх можна виписати

$$\begin{array}{llll} \omega_1\omega_4\omega_4\omega_4 & \omega_4\omega_1\omega_4\omega_4 & \omega_4\omega_4\omega_1\omega_4 & \omega_4\omega_4\omega_4\omega_1 \\ \omega_2\omega_4\omega_4\omega_4 & \omega_4\omega_2\omega_4\omega_4 & \omega_4\omega_4\omega_2\omega_4 & \omega_4\omega_4\omega_4\omega_2 \\ \omega_3\omega_4\omega_4\omega_4 & \omega_4\omega_3\omega_4\omega_4 & \omega_4\omega_4\omega_3\omega_4 & \omega_4\omega_4\omega_4\omega_3 \end{array}$$

Оскільки розподіл вважається рівномірним, то ймовірність сприятливої події A дорівнює

$$\Pr(A) = \frac{m}{n} = \frac{12}{256} = \frac{3}{64}.$$

Приклад 2.4 Знайдемо ймовірність того, що довільно вибрані з колоди 5 карт можна розташувати підряд за зростом, — наприклад, 8,9,10, В,Д, або 2,3,4,5,6.

Елементарні події тут — вибір 5 карт з колоди, що містить 52 карти (ми до мовились, що в прикладах використовуємо бріджеву колоду). Отже елементарних подій $n = C_{52}^5$.

Розподіл вважаємо рівномірним.

Сприятливі події ті, в яких вибрані карти можна розташувати підряд за зростом. Підрахуємо їх кількість.

Всього зростів 13. Найменший зріст карт в сприятливій події може бути одним із $13 - 4 = 9$ (валет, дама, король та туз найменшими в послідовності із 5 зростів бути не можуть). Отже найменший зріст може бути один із 9. Коли найменший вибрано, тоді решта зростів визначені єдиним чином — це наступні 4 зрости.

Коли зріст карт вибрано, тоді кожну карту в своєму зрості можна вибрати 4-ма способами (одну масть із 4-х потібно вибрати). Отже коли найменший зріст вибрано, тоді карти можна вибрати 4^5 способами. Отже всього сприятливих подій $m = 9 \cdot 4^5$. Проведені обчислення дозволяють знайти шукану ймовірність p — нею буде

$$p = \frac{m}{n} = \frac{9 \cdot 4^5}{C_{52}^5} = \frac{9 \cdot 4^5 \cdot 5! \cdot 47!}{52!} = \frac{192}{54145} \sim 0.0035.$$

Приклад 2.5 Знайдемо ймовірність p то, що серед 5 довільно вибраних карт виявиться рівно два тузи.

Елементарними подіями є вибір 5 карт із колоди, що містить 52 карти. Тому елементарних подій $n = C_{52}^5$.

Щоб подія була сприятливою, потрібно серед 4 тузів вибрати 2 (це можна зробити $C_4^2 = 6$ способами), а серед не-тузів потрібно вибрати 3 карти (це можна зробити C_{47}^3). Отже сприятливих подій $m = 6 \cdot C_{47}^3$.

Розподіл вважаємо рівномірним, тому

$$p = \frac{m}{n} = \frac{6 \cdot C_{47}^3}{C_{52}^5} = \frac{6 \cdot 47! \cdot 5! \cdot 47!}{3! \cdot 44! \cdot 52!} = \frac{3243}{86632} \sim 0,037.$$

Приклад 2.6 Підкидають монету і гральний кубик. Утворюємо три ймовірнісні простори: Перший — Ω_1 складається із впорядкованих пар, в яких перший елемент ω або g (цифра або герб в залежності від того, як впала монета), а другий елемент пари — число, яке випало на верхній грані кубика.

Другий ймовірнісний простір Ω_2 складається із двох подій в Ω_1 — Ω та Γ , в залежності від того, як впала монета, а третій простір Ω_3 складається із чисел $1, 2, 3, 4, 5, 6$ в залежності від того, якою виявилася верхня грань кубика.

Функцію розподілу ймовірності на Ω_1 задамо таблично див. табл. 1

	1	2	3	4	5	6
ω	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{12}$
g	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{12}$

Табл. 1: Розподіл ймовірності на прямому добутку множин $\{\omega, g\}$ та $\{1, 2, 3, 4, 5, 6\}$.

Так заповнити таблицю, як ми заповнили табл. 1, допустимо, оскільки ймовірності елементарних подій додатні і їх сума дорівнює 1.

Розподіли ймовірностей в Ω_2 та Ω_3 підраховуємо.

Які з цих трьох розподілів рівномірні?

Зрозуміло, що розподіл ймовірності на Ω_1 не рівномірний, хоча б тому, що $\Pr(\omega, 4) = \frac{1}{8} \neq \frac{1}{24} = \Pr(g, 4)$.

Подія Ω складається із 6 елементарних подій в Ω_1 :

$$\Omega = \{(\omega, 1), (\omega, 2), (\omega, 3), (\omega, 4), (\omega, 5), (\omega, 6)\},$$

Тому

$$\Pr(\Omega) = \frac{1}{12} + \frac{1}{12} + \frac{1}{12} + \frac{1}{8} + \frac{1}{24} + \frac{1}{12} = \frac{1}{2}.$$

Також

$$\Pr(\Gamma) = \frac{1}{12} + \frac{1}{12} + \frac{1}{12} + \frac{1}{8} + \frac{1}{24} + \frac{1}{12} = \frac{1}{2}.$$

Отже розподіл ймовірності на Ω_2 рівномірний.

Подібним чином переконуємося, що розподіл ймовірності на Ω_3 також рівномірний, — кожна із ймовірностей $\Pr(1), \Pr(2), \Pr(3), \Pr(4), \Pr(5), \Pr(6)$ дорівнює $\frac{1}{6}$.

2.3 Біномний розподіл, наближена формула Пуассона

Якщо послідовно n раз здійснюються випробування Бернуллі ($\Omega_1 = \{У, Н\}$ з розподілом (7)), то одержується послідовність подій $У$ та $Н$. Вводимо новий ймовірнісний простір Ω_2 , елементарними подіями в якому є послідовності випробувань Бернуллі заданої довжини.

Якщо послідовність a містить n_1 Успіхів і $n_2 = n - n_1$ Невдач, то в ймовірнісному просторі Ω_2 їй приписується ймовірність

$$\Pr(a) = p^{n_1} \cdot q^{n_2}. \quad (8)$$

Таке приписування ґрунтується на уяві про те, що результат кожного наступного випробування не залежить від того, які результати одержані перед цим, і відповідно, конкретна серія результатів має ймовірність, що дорівнює добутку ймовірностей окремих випробувань. Дещо детальніше це ми обговоримо в розділі, що стосується незалежних випробувань.

Створюється новий ймовірнісний простір Ω_3 , елементарними подіями в якому є події в Ω_2 . Точніше, $\Omega_3 = \{\omega_0, \omega_1, \omega_2, \dots, \omega_n\}$, де елементарна подія $\omega_k, 0 \leq k \leq n$ є подією в Ω_2 , яка полягає в тому, що одержана послідовність містить k успіхів.

Для прикладу, розглянемо випадок $n = 3$. Тоді

$$\Omega_2 = \{УУУ, УУН, УНУ, УНН, НУУ, НУН, ННУ, ННН\}$$

$$\Omega_3 = \{\omega_0, \omega_1, \omega_2, \omega_3\},$$

$$\omega_0 = \{\text{ННН}\}, \omega_1 = \{\text{УНН, НУН, ННУ}\}, \omega_2 = \{\text{УУН, УНУ, НУУ}\}, \omega_3 = \{\text{УУУ}\}$$

Події ω_k в ймовірнісному просторі Ω_2 складається із C_n^k елементарних подій, кожна з яких має ймовірність $p^k q^{n-k}$. Тому в новому ймовірнісному просторі Ω_3

$$\Pr(\omega_k) = b(k, n, p) = C_n^k p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n. \quad (9)$$

Розподіл (9) ймовірності в Ω_3 називають біномним розподілом.

Приклад 2.7 *В урні є 7 білих та 5 чорних куль. Ми вибираємо кулю із урни, дивимось на її колір, і повертаємо кулю в урну. І так робимо 20 разів. Яка ймовірність того, що рівно 3 рази ми побачили чорну кулю.*

Тут успіх — побачили чорну кулю, є подією в ймовірнісному просторі, де елементарними подіями є вибір однієї кулі із 12. Вважаємо розподіл ймовірностей там рівномірним і ймовірність успіху тепер обчислюється $p = \frac{5}{12}$. Відповідно $q = \frac{7}{12}$. Тепер за біномним розподілом шукана ймовірність $\Pr(\omega_3)$ обчислюється за формулою

$$\Pr(\omega_3) = C_{20}^3 p^3 q^{17} = 0.0086.$$

2.4 Розподіл Пуассона

На сьогодні практика зустрічається з дуже довгими послідовностями подій (наприклад, пам'ять в обчислювальній машині уже вимірюється терабайтами), а також повинна враховувати дуже малі ймовірності (пам'ять в обчислювальній машині надійна, ймовірність відмови є малою величиною). За таких умов обчислення $b(k, n, p)$ за формулою (9) може викликати проблеми. Сімеон-Дені Пуассон запропонував наближену

формулу для $b(k, n, p)$, коли n велике, p мале, а $\lambda = np$ і k знаходяться в розумних межах:

$$b(k, n, p) \sim \frac{\lambda^k}{k!} e^{-\lambda} = p(k, \lambda).$$

Перед тим, як обґрунтовувати таке наближення нагадаємо визначення числа Непера e :

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Відповідно,

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda k}{n}\right)^n = \lim_{n \rightarrow \infty} \left(\left(1 - \frac{\lambda k}{n}\right)^{-\frac{n}{\lambda k}}\right)^{-\lambda k} = e^{-\lambda k}. \quad (10)$$

Обґрунтовується наближена формула для $b(k, n, p)$ наступним чином. Для заданих n, p і, відповідно, $\lambda = n \cdot p$, де n велике, p мале, а числа k, λ лежать в розумних межах виписуємо послідовність

$$b(0, n, p) \sim p(0, n), \quad b(1, n, p), \dots, b(k-1, n, p), \quad b(k, n, p) \sim p(k, n), \dots$$

Далі робота з обґрунтування розбивається на 4 частини:

- 1) Обчислюємо перший елемент $b(0, n, p)$;
- 2) Шукаємо рекурентне співвідношення для послідовності. Для цього обчислюємо відношення наступного елемента послідовності до попереднього:

$$\frac{b(k, n, p)}{b(k-1, n, p)}.$$

- 3) Обчислюємо кілька перших елементів послідовності і відгадуємо формулу для загального члена.

- 4) Доводимо формулу для загального члена послідовності, яка має відомий перший член і відому рекурентну формулу, методом математичної індукції.

Переходимо до частини 1), вважаючи $\lambda = np$, $p = \frac{\lambda}{n}$.

$$b(0, n, p) = C_n^0 p^0 q^n = q^n = (1 - p)^n = \left(1 - \frac{\lambda}{n}\right)^n \sim e^{-\lambda}.$$

Переходимо до частини 2).

$$\begin{aligned} \frac{b(k, n, p)}{b(k-1, n, p)} &= \frac{C_n^k p^k q^{n-k}}{C_n^{k-1} p^{k-1} q^{n-k+1}} = \frac{n!}{k!(n-k)!} \cdot \frac{(k-1)!(n-k+1)!}{n!} \cdot \frac{p^k q^{n-k}}{p^{k-1} q^{n-k+1}} \\ &= \frac{(n-k+1)p}{kq} = \frac{\lambda - kp + p}{kq}. \end{aligned}$$

В останньому виразі беремо $p = 0$, а $q = 1$, одержуємо рекурентне співвідношення

$$\frac{b(k, n, p)}{b(k-1, n, p)} \sim \frac{p(k, \lambda)}{p(k-1, \lambda)} = \frac{\lambda}{k}$$

Переходимо до частини 3 — виписуємо перші члени послідовності $p(k, \lambda)$:

$$p(0, \lambda) = e^{-\lambda}, p(1, \lambda) = \frac{\lambda^1}{1!}, p(2, \lambda) = \frac{\lambda^2}{2!}, p(3, \lambda) = \frac{\lambda^3}{3!}, \dots$$

і вгадану загальну формулу

$$p(k, \lambda) = \frac{\lambda^k}{k!}.$$

можна довести метою повної математичної індукції, що є четвертою частиною доведення.

Якщо простір подій Ω нескінченний, $\Omega = \{\omega_0, \omega_1, \omega_2, \dots\}$, то розподіл

$$\Pr(\omega_k) = p(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots \quad (11)$$

називається розподілом Пуассона з параметром λ .

Розподіл Пуассона введений коректно, тому що числа $\frac{\lambda^k}{k!}e^{-\lambda}$ додатні при $k \geq 0$, і

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

Якщо ж ймовірнісний простір Ω скінченний, то формули (11) задають прийнятне наближення біномного розподілу.

Приклад 2.8 *Яка ймовірність того, що в групі із 25 студентів бодай у одного день народження прийдеться на день іспиту з дискретної математики.*

Ймовірність, яку потрібно знайти, позначимо через x . Отже, x це ймовірність того, що бодай у одного студента ден народження припадає на день іспиту з дискретної математики. Протилежною подією буде те, що у жодного студента день народження не збігається із днем іспиту. Ймовірність цієї протилежної події позначимо через y . Таким чином,

$$x + y = 1.$$

Випробуванням є перевірка дня народження певного студента — якщо вона збігається з днем іспиту, то маємо Успіх, а якщо не збігається, то маємо Невдачу.

Будемо вважати, ймовірність Успіху дорівнює

$$p = \frac{1}{365}.$$

Відповідно, ймовірність Невдачі дорівнює

$$q = \frac{364}{365}.$$

Будемо шукати ймовірність y за формулою Пуассона. Для цього випишемо відповідні параметри

$n = 25$ — кількість випробувань;

$\lambda = np = \frac{25}{365}$ — параметр розподілу Пуассона;

$k = 0$ — кількість Успіхів в серії випробувань.

За формулою Пуассона

$$y = p(0, \lambda) = \frac{\left(\frac{25}{365}\right)^0}{0!} e^{-\frac{25}{365}} = e^{-\frac{25}{365}} = 0.9337998560... \sim 0.93.$$

За біномним розподілом

$$y = C_{25}^0 p^0 \cdot q^{25} = \left(\frac{364}{365}\right)^{25} = 0.9466084651... \sim 0.95$$

Відповідно, з використанням розподілу Пуассона маємо

$$x = 1 - y \sim 0.07,$$

а з використанням біномного розподілу маємо

$$x = 1 - y \sim 0.05.$$

Помітна різниця зумовлена малою довжиною випробувань.

2.5 Гіпергеометричний розподіл

Є так званий гіпергеометричний ряд (ступеневий ряд по z):

$$F(\alpha, \beta, \gamma, z) = 1 + \frac{\alpha\beta}{\gamma} \cdot \frac{z}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)} \cdot \frac{z^2}{2!} + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{\gamma(\gamma+1)(\gamma+2)}$$

Коефіцієнт при $z^n, n \geq 2$ дорівнює

$$\frac{\alpha(\alpha+1)(\alpha+2) \dots (\alpha+n-1)\beta(\beta+1)(\beta+2) \dots (\beta+n-1)}{\gamma(\gamma+1)(\gamma+2) \dots (\gamma+n-1)} \cdot \frac{1}{n!} \quad (12)$$

Коли α, β, γ є натуральними числами, тоді вираз 12 можна переписати наступним чином

$$\frac{\alpha(\alpha+1)\dots(\alpha+n-1)}{n!} \cdot \frac{\beta(\beta+1)\dots(\beta+n-1)}{n!} \cdot \frac{n!}{\gamma(\gamma+1)\dots(\gamma+n-1)} \quad (13)$$

або у вигляді

$$\frac{C_{\alpha+n-1}^n \cdot C_{\beta+n-1}^n}{C_{\gamma+n-1}^n} \quad (14)$$

Тепер розглянемо задачу, що приводить до використання гіпергеометричного розподілу.

Приклад 2.9 *Маємо n куль. Із них n_1 червона, а n_2 чорна, $n_1 + n_2 = n$. Із цих n куль вибираємо r , $0 \leq r \leq n$ куль. Потрібно знайти ймовірність того, що буде вибрано рівно k , $0 \leq k \leq n_1$ червоних куль.*

Конкретний вибір r куль це елементарна подія. Кількість усіх елементарних подій — це кількість r —елементних підмножин у множині із r елементів, тобто C_n^r .

Сприятлива подія — вибір k червоних куль і вибір $r - k$ чорних куль. Червоні кулі можна вибрати $C_{n_1}^k$, а чорні можна вибрати $C_{n-n_1}^{r-k}$ способами. Отже за правилом добутку, сприятливий подій — $C_{n_1}^k \cdot C_{n-n_1}^{r-k}$.

Ймовірність q_k сприятливої події обчислюється за формулою

$$q_k = \frac{C_{n_1}^k \cdot C_{n-n_1}^{r-k}}{C_n^r} \quad (15)$$

Розподіл, при якому ймовірність q_k величини k обчислюється за формулою (15), називається гіпергеометричним, а числа n, n_1, r , що використовуються в (15), називаються параметрами гіпергеометричного розподілу.

Приклад 2.10 *Нехай в партії із 100 деталей є 10 бракованих. Яка ймовірність того, що серед вибраних 15 деталей виявиться рівно 2 браковані.*

Розв’язування. Вього елементарних подій (вибору 15 деталей із 100) є C_{100}^{15} . Дві браковані деталі можна вибрати C_{10}^2 , а 13 небракованих можна вибрати C_{90}^{13} . Отже всього сприятливих випадків буде

$$C_{10}^2 \cdot C_{90}^{13}.$$

За формулою (15) обчислюємо потрібну ймовірність x :

$$x = \frac{C_{10}^2 \cdot C_{90}^{13}}{C_{100}^{15}} = 0.2919112282 \dots \sim 0.29.$$

В наступних двох прикладах оцінимо параметри гіпергеометричного розподілу. Такі задачі розв’язуються в матстатистиці.

Приклад 2.11 *Якою найімовірніше є кількість бракованих деталей в партії із 100, коли серед вибраних 15 їх виявилось 2.*

Один із способів — припустити, що вибір бракованої деталі серед 15 вибраних має ту ж ймовірність, що і вибір бракованої деталі серед усіх 100. Тому

$$\frac{2}{15} = \frac{x}{100}, \quad 15x = 200, \quad x \in \{13, 14\}.$$

Другий підхід полягає в прямому обчисленні ймовірності $y(x)$ того, що серед вибраних 15 деталей буде 2 браковані, якщо вся партія із 100 деталей містить $x = 2, 3, \dots, 87$ бракованих деталей. $y(x)$ обчислюємо за формулою

$$y(x) = \frac{C_x^2 \cdot C_{100-x}^{13}}{C_{100}^{15}}$$

В нашому прикладі

$$\dots, y(10) \sim 0.291, \quad y(11) \sim 0.305, \quad y(12) \sim 0.312, \\ y(13) \sim 0.315, \quad y(14) \sim 0.312, \dots$$

Послідовність $y(x)$ має найбільший елемент $y(13)$.
Отже відповіддю буде 13.

Приклад 2.12 *Якою найімовірніше є кількість деталей в партії, якщо бракованих в ній 100, а у вибірці 15 деталей виявилось бракованих дві.*

Як і вище, ми можемо припустити, що ймовірність вибору бракованої серед вибраних 100 буде така ж, як і ймовірність вибору бракованої деталі у всій партії, що містить x деталей. Тому

$$\frac{2}{100} = \frac{100}{x}, \quad 2x = 10000, \quad x = 5000.$$

Знову ж, можна побудувати послідовність ймовірностей $z(n)$ того, що серед 100 обраних деталей в партії із n деталей виявиться рівно 2 браковані. Обчислення показують:

$$z(4998) = 0.2761850365, \quad z(4999) = 0.2761850596, \\ z(5000) = 0.2761850596, \quad z(5001) = 0.2761850366.$$

що з точністю до 10 знаків після коми $z(4999) = z(5000)$, а на решті значень n $z(n)$ приймає менші значення.

2.6 Геометричний розподіл

Коли елементарні події розташовані у послідовність $\Omega = \{\omega_0, \omega_1, \omega_2, \dots, \omega_n, \dots\}$ і послідовність ймовірностей елементарних подій утворює геометричну прогресію,

$$\Pr(\omega_0) = q, \quad \Pr(\omega_1) = pq, \dots, \Pr(\omega_n) = p^n q, \dots \quad 0 < p, q < 1, \quad p+q=1$$

то кажуть, що маємо геометричний розподіл

Приклад 2.13 Нехай маємо випробування Бернуллі $\{Y, H\}$ з розподілом Бернуллі $\Pr(Y) = p, \quad \Pr(H) = q$. Проводиться серія таких випробувань до тих пір, поки не буде одержаний Успіх. Яка ймовірність того, що Успіх вперше буде досягнуто в десятому випробуванні.

Конкретна послідовність випробувань має ймовірність $p^{n_1}q^{n_2}$, де n_1 кількість Успіхів в послідовності (серії), а n_2 — кількість Невдач. Тому на послідовностях випробувань, що закінчуються першим Успіхом, виникає розподіл:

ω	$\Pr(\omega)$
$\omega_0 = \{Y\}$	$\Pr(\omega_0) = p$
$\omega_1 = \{HY\}$	$\Pr(\omega_0) = pq$
$\omega_1 = \{HHY\}$	$\Pr(\omega_0) = pq^2$
$\dots$	$\dots$
$\omega_n = \{\underbrace{HH \dots H}_{n-1} Y\}$	$\Pr(\omega_n) = pq^n$
$\dots$	$\dots$

Цей розподіл є геометричним, і потрібна ймовірність дорівнює pq^9 .

2.7 Від'ємний біномний розподіл

Приклад 2.14 Проводиться серія випробувань Бернуллі з розподілом Бернуллі $\Pr(Y) = p, \Pr(H) = q, 0 < p, q < 1$. Випробування припиняються, коли настає $k+1$ -й Успіх, $k = 0, 1, 2, \dots$. Знайдемо ймовірність $\Pr(\omega_n)$ елементарної події ω_n , яка полягає в тому, що $k+1$ -й Успіх станеться на $n+k+1$ -му місці, $n = 0, 1, 2, \dots$, — при $n+k+1$ -му випробуванні.

Конкретні послідовності Успіхів та Невдач у відповідному просторі подій мають ймовірність $p^{n_1} \cdot q^{n_2}$, де n_1 — кількість Успіхів у послідовності, а n_2 — кількість Невдач. Тому послідовність довжини $n + k + 1$, що містить $k + 1$ Успіх, має ймовірність $p^{k+1}q^n$. Послідовностей, що мають $k + 1$ -й Успіх на $n + k + 1$ -му місці, C_{n+k}^n : щоб одержати $k + 1$ Успіхів в послідовності довжини $n + r$ і при цьому на останньому місці мати Успіх, потрібно решту k Успіхів розташувати на попередніх $n + r - 1$ -му місці, а $C_{n+k}^k = C_{n+k}^n$.

Отже, оскільки

- ймовірність послідовності Успіхів та Невдач довжини $n + k + 1$, що містить $k + 1$ Успіх і на останньому місці має Успіх, дорівнює $p^{k+1}q^n$;
- кількість послідовностей Успіхів та Невдач довжини $n + k + 1$, що містить $k + 1$ Успіх і на останньому місці мають Успіх, дорівнює C_{n+k}^k ,

то ймовірність події ω_n дорівнює $C_{n+k}^k p^{k+1} q^k$:

$$\Pr(\omega_n) = C_{n+k}^k p^{k+1} q^k. \quad (16)$$

Перевіримо, що

$$\sum_{n=0}^{\infty} \Pr(\omega_n) = \sum_{n=0}^{\infty} C_{n+k}^k p^{k+1} q^k = 1. \quad (17)$$

Для цього використаємо породжуючу функцію послідовності, яку уже використовували в комбінаториці і детальніше зупиняємося на стор. 82 — виписуємо суму членів нескінченної спадної геометричної прогресії із першим членом 1 і знаменником qz , $0 < qz < 1$:

$$f(z) = \frac{1}{1 - qz} = 1 + (qz) + (qz)^2 + \dots = \sum_{n=0}^{\infty} q^n z^n.$$

Знаходимо похідні по z від лівої та правої частини одержаної рівності:

$$\frac{d}{dz}f(z) = \frac{q}{(1 - qz)^2} = q + 2q^2z + 3q^3z^2 + \dots + (n+1)q^{n+1}z^n + \dots = \sum_{n=0}^{\infty} (n+1)q^{n+1}z^n$$

$$\frac{d^2}{dz^2}f(z) = \frac{2 \cdot 1 \cdot q^2}{(1 - qz)^3} = 2 \cdot 1 \cdot q^2 + 3 \cdot 2 \cdot q^3z + \dots = \sum_{n=0}^{\infty} (n+2)(n+1)q^{n+2}z^n.$$

$$\frac{d^k}{dz^k}f(z) = \frac{k!q^k}{(1 - qz)^{k+1}} = \frac{k!}{0!}q^k + \frac{(k+1)!}{1!}q^{k+1}z + \dots = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!}q^{n+k}z^n.$$

Одержуємо при $k = 0, 1, 2, \dots$

$$\frac{q^k}{(1 - qz)^{k+1}} = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!k!}q^{n+k}z^n.$$

Розділивши останню рівність на q^k і домноживши на p^{k+1} Одержуємо рівність

$$\frac{p^{k+1}}{(1 - qz)^{k+1}} = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!k!}q^n p^{k+1}z^n = \sum_{n=0}^{\infty} C_{n+k}^k q^n p^{k+1}z^n. \quad (18)$$

В рівність (18) підставляємо 1 замість z і врахувавши, що $1 - q = p$, одержуємо потрібну нам рівність (17)

Законність виконаних нами дій з нескінченними сумами забезпечується тим, що $0 < p \cdot q < 1$ і обгрунтовується методами математичного аналізу. Заглиблюватися в цю проблематику не будемо. Більш детально можливості роботи з нескінченними сумами обговорюється в розділі, що стосується породжуючих функцій, формальних степеневих рядів.

Рівність (17) дозволяє дати наступне означення

Розподіл ймовірності

$$\Pr(\omega_n) = C_{n+r}^r p^{r+1} q^n, \quad (19)$$

де $r = 0, 1, 2, \dots$ — параметр розподілу, і $n = 0, 1, 2, \dots$, на просторі подій $\omega_0, \omega_1, \dots$ називається від’ємним біномним.

Які міркування використовуються при введенні терміну “від’ємний біномний”?

Згадаємо, що біномний коефіцієнт C_x^r підраховується за формулою:

$$C_x^r = \frac{x \cdot (x-1) \cdot \dots \cdot (x-r+1)}{1 \cdot 2 \cdot \dots \cdot (r-1) \cdot r} \quad (20)$$

В цій формулі і чисельник і знаменник є добутком r множників, тому в ній обов’язково передбачається, що r — натуральне число або 0. Добуток порожньої сукупності множників дорівнює одиниці. Тому, коли $r = 0$, тоді і чисельник і знаменник правої частини мають порожню сукупність множників. Отже в цьому випадку і чисельник і знаменник дорівнюють 1. Сказане дозволяє стверджувати, що права частина рівності (20) існує при будь-якому дійсному x . Відмітимо значення C_x^r при деяких значеннях x, r :

$$C_0^0 = 1, \quad C_0^r = 0 \text{ коли } r > 0, \quad C_{r-1}^r = 0 \text{ коли } r > 0.$$

Зупинимося на випадку, коли у формулі (20) замість x підставити $-n$. Маємо

$$\begin{aligned} C_{-n}^r &= \frac{(-n) \cdot (-n-1) \cdot \dots \cdot (-n-r+1)}{1 \cdot 2 \cdot \dots \cdot (r-1) \cdot r} = \\ &= (-1)^r \frac{(n) \cdot (n+1) \cdot \dots \cdot (n+r-1)}{1 \cdot 2 \cdot \dots \cdot (r-1) \cdot r} = (-1)^r \frac{(n+r-1) \cdot (n+r-2) \cdot \dots \cdot n}{1 \cdot 2 \cdot \dots \cdot (r-1) \cdot r} \end{aligned}$$

$$C_{-n}^r = (-1)^r C_{n+r-1}^r. \quad (21)$$

Формула (21) і є мотивацією введення терміну “від’ємний біномний”.

2.8 Поліномний розподіл

Нехай

$$(z_1 + z_2 + \dots z_n)^m = \sum_{i_1+i_2+\dots+i_n=m} a_{i_1 i_2 \dots i_n} z_1^{i_1} z_2^{i_2} z_3^{i_3} \dots z_n^{i_n} \quad (22)$$

При підраховуванні коефіцієнта $a_{i_1 i_2 \dots i_n}$ міркують таким чином. Розкривши дужки в лівій частині (22) ми одержимо суму добутків змінних $z_1, \dots, z_n$ взятих по одній із кожного множника, із кожної із n дужок. Щоб одержати $z_1^{i_1}$ потрібно із n дужок вибрати i_1 і в цих дужках вибрати z_1 . Це можна зробити $C_n^{i_1}$ способами. Із решти $n - i_1$ дужок ми повинні вибрати i_2 дужок (це можна зробити $C_{n-i_1}^{i_2}$ способами) і в цих дужках вибрати z_2 . Продовжуючи цей процес ми знайдемо

$$a_{i_1 i_2 \dots i_n} = C_n^{i_1} \cdot C_{n-i_1}^{i_2} \cdot C_{n-i_1-i_2}^{i_3} \cdot C_{i_n}^{i_n} = \frac{n!}{i_1!(n-i_1)!} \cdot \frac{(n-i_1)!}{(n-i_1-i_2)!i_2!} \cdot \dots \cdot \frac{i_n!}{i_n!}.$$

Після скорочення одержуємо

$$a_{i_1 i_2 \dots i_n} = \frac{n!}{i_1! i_2! \dots i_n!} \quad (23)$$

Коефіцієнт $a_{i_1 i_2 \dots i_n}$ в (22), який обчислюється за формулою (23), називається поліномним коефіцієнтом. Сказане є мотивацією введення словосполучки “поліномний розподіл”

Нехай певне випробування може закінчитися n результатами: $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ із імовірностями

$$\Pr(\omega_i) = p_i, \quad 0 \leq p_i \leq 1, (i = 1, 2, \dots, n), \sum_{i=1}^n p_i = 1.$$

Для прикладу, підкидання гральної кістки може закінчитися випадінням однієї грані із 6 з ймовірністю $\frac{1}{6}$.

Розглядаємо новий простір подій Ω_1 , в якому елементарними подіями є послідовності заданої довжини m елементарних подій із попереднього простору $\omega \in \Omega$. Цим новим елементарним подіям (послідовностям) приписується ймовірність — добуток ймовірностей тих $\omega \in \Omega$, що входять у послідовність. Так, коли $m = 6$, $\omega_1, \omega_2, \omega_3 \in \Omega$, і

$$a = (\omega_3 \omega_3 \omega_2 \omega_1 \omega_2 \omega_3) \in \Omega_1,$$

то

$$\Pr(a) = (\Pr(\omega_1)) \cdot (\Pr(\omega_2))^2 \cdot (\Pr(\omega_3))^3.$$

В просторі подій Ω_1 розглядаємо події $A_{i_1 i_2 \dots i_n}$, що складаються із послідовностей $a \in \Omega_2$, що мають i_1 входжень ω_1 , i_2 входжень ω_2 , i_3 входжень ω_3 , ... , i_n входжень ω_n $i_1 + i_2 + \dots + i_n = m$. За домовленістю, ймовірність кожної елементарної події, що входить в $A_{i_1 i_2 \dots i_n}$, дорівнює $p_1^{i_1} p_2^{i_2} p_3^{i_3} \dots p_n^{i_n}$. А в $A_{i_1 i_2 \dots i_n}$ міститься $\frac{n!}{i_1! i_2! \dots i_n!}$ елементарних подій із Ω_1 . Тому

$$\Pr(A_{i_1 i_2 \dots i_n}) = \frac{n!}{i_1! i_2! \dots i_n!} p_1^{i_1} p_2^{i_2} p_3^{i_3} \dots p_n^{i_n}. \quad (24)$$

Розподіл (24) в просторі подій Ω_2 , елементарними подіями в якому є $A_{i_1 i_2 \dots i_n}$, називається поліномним.

Приклад 2.15 Яка ймовірність того, що при підкиданні 12 гральних кісток кожна грань випаде двічі?

В цьому прикладі $p_1 = p_2 = p_3 = p_4 = p_6 = \frac{1}{6}$ і шукана ймовірність дорівнює

$$\frac{12!}{2!2!2!2!2!2!} \cdot \left(\frac{1}{6}\right)^{12} = \frac{479001600}{64} \frac{1}{2176782336} = \frac{1925}{559872} = 0,0034...$$

2.9 Статистики Максвелла — Больцмана, Бозе — Енштейна, Фермі — Дирака

Розглядаємо простір, що розділений на $n \geq 1$ комірок, в якому містяться $r \geq 1$ частинок. В кожен момент часу кожна частинка знаходиться в певній комірці — між комірками знаходитися не може. Випадкова подія $i = (i_1, i_2, \dots, i_n)$ полягає в тому, що вимірювання показує певне заповнення комірок — i_1 частинок в першій комірці, i_2 частинок в другій комірці і т.д. Зрозуміло, що повинна виконуватись рівність

$$i_1 + i_2 + \dots + i_n = r. \quad (25)$$

Простір подій Ω містить стільки елементарних подій, скільки невід'ємних цілих розв'язків у рівніння (25), точбто стільки ж, скільки є сполук з повтореннями із n елементів по r :

$$|\Omega| = C_{n+r-1}^r.$$

Для деяких частинок всі елементарні подій $\omega \in \Omega$ рівноймовірні:

$$\text{Pr}(\omega) = (C_{n+r-1}^r)^{-1}.$$

В такому випадку кажуть, що частинки “підкоряються статистиці Бозе — Енштейна”.

Якщо ж розподіл ймовірності на Ω поліномний, тобто подія $i = (i_1, i_2, \dots, i_n)$ має ймовірність

$$\text{Pr}(i) = \frac{r!}{i_1! i_2! \dots i_n!} \left(\frac{1}{n} \right)^r,$$

то кажуть, що частинки “підкоряються статистиці Максвелла — Больцмана”.

Для частинок третього типу виконується додаткова умова — в одній комірці не може знаходитися більше

однієї частинки. Для таких частинок елементарних подій з ненульовою ймовірністю C_n^r . Досліди показують, що всі такі елементарні події можна вважати рівноймовірними. Отже кожна елементарна подія $i = (i_1, i_2, \dots, i_n)$, що підкоряється умовам

$$i_1 + i_2 + \dots + i_n = r, \quad 0 \leq i_1, i_2, \dots, i_n \leq 1,$$

має ймовірність

$$\Pr(i) = (C_n^r)^{-1}.$$

Про частинки, які дають такий розподіл ймовірності, кажуть, що вони “підкоряються статистиці Фермі — Дирака.”

2.10 Локальна та інтегральна теореми Лапласа

Розглянемо ще одне наближення біномного розподілу — формули Бернуллі $b(k, n, p) = C_n^k p^k q^{n-k}$, — одне, формулу Пуасона, ми вже розглянули.

Введемо функцію

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

— її графік зображений на рис. 2.10.

Теорема 2.1 (Локальна теорема Лапласа) *Розглядаються послідовності довжини n випробувань Бернуллі з розподілом Бернуллі $\Pr(Y) = p$, $\Pr(H) = q$. Вважаємо, що коли в послідовності n_1 раз зустрічається Успіх і n_2 разів зустрічається Невдача, тоді ймовірність появи такої послідовності дорівнює $p^{n_1} q^{n_2}$. За таких умов ймовірність $b(k, n, p)$ появи рівно k успіхів в послідовності довжини n асимптотично дорівнює*

$$\frac{1}{\sqrt{npq}} \varphi\left(\frac{k - np}{\sqrt{npq}}\right)$$

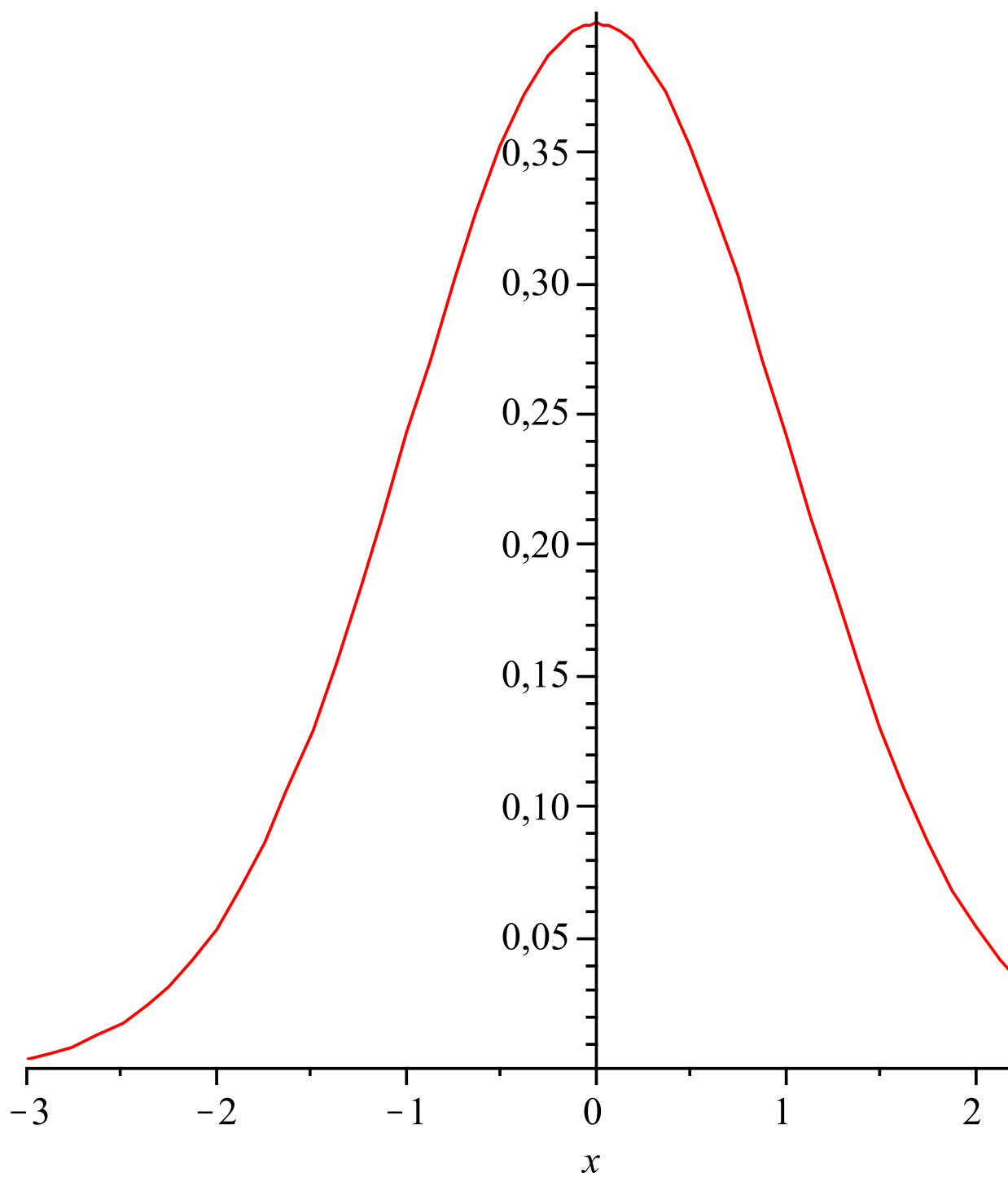


Рис. 2: Графік функції $\phi(x)$

в тому розумінні, що

$$\lim_{n \rightarrow \infty} \frac{b(k, n, p)}{\frac{1}{\sqrt{npq}} \varphi\left(\frac{k - np}{\sqrt{npq}}\right)} = 1.$$

Приклад 2.16 Знайти ймовірність a того, що Успіх наступить рівно 70 разів в 400 випробуваннях Бернулі, якщо в одному випробуванні ймовірність успіху дорівнює 0,3.

При знаходженні потрібної ймовірності a використаємо локальну теорему Лапласа з параметрами

$$n = 400, k = 70, p = 0.3.$$

Проводимо обчислення:

$$k - np = -50, \sqrt{npq} = \sqrt{8400} = 91.65151390,$$

$$x = \frac{k - np}{\sqrt{npq}} = -0.545 \dots, \quad \varphi(x) = 0.3437 \dots$$

Відповіддю буде число

$$a = \frac{\varphi(x)}{\sqrt{npq}} = \frac{0.3437818233}{91.65151390} = 0.004 \dots$$

Теорема 2.2 (Інтегральна теорема Лапласа) За тих же умов, що і в локальній теоремі, ймовірність a того, що в серії випробувань довжини n Успіх настане k разів, $0 \leq k_1 \leq k \leq k_2$, асимптотично дорівнює

$$b = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{z^2}{2}} dz$$

де

$$x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

“Асимптотично дорівнює” означає, що

$$\lim_{n \rightarrow \infty} \frac{a}{b} = 1.$$

Доведення і локальної і інтегральної теорем Лапласа виходить за межі дискретної теорії ймовірності і ми наводити його не будемо.

При розв’язуванні задач, які вимагають використання інтегральної теореми Лапласа, користуються таблицями значень функції

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

Приклад 2.17 Знайти ймовірність а того, що Успіх наступить від $k_1 = 70$ до $k_2 = 140$ разів в $n = 400$ випробуваннях Бернуллі, якщо в одному випробуванні ймовірність успіху дорівнює 0,3.

Як і в прикладі 2.16, маємо

$$n = 400, k_1 = 70, p = 0.3, k_1 - np = -50, k_2 = 140,$$

$$k_2 - np = 20, \sqrt{npq} = \sqrt{8400} = 91.65 \dots,$$

$$x_1 = \frac{k_1 - np}{\sqrt{8400}} = -0.545 \dots, \quad x_2 = \frac{k_2 - np}{\sqrt{8400}} = 0.218 \dots$$

$$\int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx = 0.736 \dots$$

Відповіддю буде число b :

$$b = \frac{0.736 \dots}{\sqrt{2\pi}} = 0.293 \dots$$

3 Умовна ймовірність

3.1 Незалежні події

Щоб підкреслити незалежність певних подій, вживають приказку “на городі бузина, а в Києві дядько”. Формалізуємо ймовірнісний простір (тобто простір подій разом з розподілом ймовірності в ному), в якому розглядається випадкова подія “на городі бузина, а в Києві дядько”, і введемо поняття “незалежні події”.

Маємо дві групи студентів — в першій $n_1 = 20$, а в другій $n_2 = 23$ студенти. В першій із них $m_1 = 4$ студенти мають на городі бузину, а в другій $m_2 = 7$ студентів мають в Києві дядька. Виникає два простори подій — Ω_1 , елементарними подіями в якому є вибір одного студента з першої групи з рівномірним розподілом, і Ω_2 , елементарними подіями в якому є вибір одного студента з другої групи з рівномірним розподілом. В першому просторі подій є подія A — вона складається із тих елементарних подій, при яких вибраний студент має на городі бузину. В другому просторі подій є подія B — вона складається із тих елементарних подій, при яких вибраний студент має в Києві дядька. Оскільки розподіли рівномірні, то

$$\Pr(A) = \frac{m_1}{n_1} = \frac{4}{20} = \frac{1}{5}, \quad \Pr(B) = \frac{m_2}{n_2} = \frac{7}{23}.$$

Розглянемо новий простір подій — $\Omega = \Omega_1 \times \Omega_2$. Нагадаємо, що прямим добутком двох множин називається множина впорядкованих пар, в яких перший елемент береться із першої множини, а другий — із другої множини. За правилом добутку

$$|\Omega| = |\Omega_1| \cdot |\Omega_2| = n_1 \cdot n_2 = 20 \cdot 23 = 460.$$

Елементарною подією в Ω є вибір пари студентів —

одного із першої групи, а другого із другої групи. Розподіл ймовірності в Ω вважаємо рівномірним.

В новому просторі елементарними подіями є вибір двох студентів — один із першої групи і другий із другої групи. Також в новому просторі маємо три події

- A_1 — “на городі бузина” — $A \times \Omega_2$;
- B_1 — “в Києві дядько” — $\Omega_1 \times B$;
- $A_1 B_1$ — на “городі бузина а в Києві дядько” — $A \times B$

Підрахуємо ймовірності цих подій

$$\begin{aligned}\Pr(A_1) &= \frac{m_1 n_2}{n_1 n_2} = \frac{m_1}{n_1} = \frac{1}{5}; \\ \Pr(B_1) &= \frac{n_1 m_2}{n_1 n_2} = \frac{m_2}{n_2} \\ \Pr(A_1 B_1) &= \frac{m_1 m_2}{n_1 n_2} = \Pr(A_1) \cdot \Pr(B_1).\end{aligned}$$

Визначальною умовою незалежності двох подій на одному ймовірнісному просторі є те, що ймовірність їх перетину дорівнює добутку ймовірності першої події на ймовірність другої події.

Приклад 3.1 *В групі із 20 студентів 5 студентів мають на городі бузину, і 5 мають в Києві дядька. Студентів, що мають на городі бузину і мають в Києві дядька немає. Чи є незалежними події A — студент має на городі бузину, і B — студент має в Києві дядька*

Відповідь — ні події не є незалежними (є залежними), тому що подія A має ненульову ймовірність — $\frac{1}{4}$, і подія B має ненульову ймовірність — $\frac{1}{4}$, а подія $A \cap B$ має нульову ймовірність, — $\Pr(A) \cdot \Pr(B) \neq \Pr(A \cap B)$.

Приклад 3.2 *Група студентів складається із 36 осіб. Із них 12 мають на городі бузину і 12 мають в Києві дядька. Скільки повинно бути тих студентів, які*

і мають бузину і мають дядька, щоб дві події A — “мати на городі бузину” і B — “мати в Києві дядька” були незалежні.

Для одержання відповіді позначимо шукану кількість через x і підрахуємо ймовірності

$$\Pr(A) = \frac{12}{36} = \frac{1}{3} = \Pr(B); \quad \frac{x}{36} = \Pr(AB).$$

Щоб події були незалежними, потрібно, щоб виконувалася рівність

$$\frac{1}{3} \cdot \frac{1}{3} = \frac{x}{36},$$

Тобто $x = 4$.

Визначення 3.1 *Події A, B в певному ймовірнісному просторі вважаються незалежними, якщо $\Pr(AB) = \Pr(A) \cdot \Pr(B)$.*

Приклад 3.3 *Розглянемо простір подій із прикладу 2.6 з розподілом, що заданий таблицею 1. Чи є події Π та 1, Π та 4, Γ та 5 незалежними.*

В прикладі 2.6 проведені обчислення, що дали результати

$$\Pr(\Pi) = \Pr(\Gamma) = \frac{1}{2}, \quad \Pr(1) = \Pr(4) = \Pr(5) = \frac{1}{6}.$$

Ймовірності добутоків подій знаходимо із таблиці 1:

$$\Pr(\Pi \cdot 1) = \Pr(\pi, 1) = \frac{1}{12},$$

$$\Pr(\Pi \cdot 4) = \Pr(\pi, 4) = \frac{1}{8},$$

$$\Pr(\Gamma \cdot 5) = \Pr(\gamma, 4) = \frac{1}{8}.$$

Оскільки

$$\Pr(\Pi \cdot 1) = \frac{1}{2} \cdot \frac{1}{6} = \Pr(\Pi) \cdot \Pr(1),$$

$$\Pr(\Pi \cdot 4) \neq \frac{1}{2} \cdot \frac{1}{6} = \Pr(\Pi) \cdot \Pr(4),$$

$$\Pr(\Gamma \cdot 5) \neq \frac{1}{2} \cdot \frac{1}{6} = \Pr(\Gamma) \cdot \Pr(5),$$

то події Π та 1 незалежні, а подій Π та 4, і Γ та 5 не є незалежними

3.2 Ймовірність послідовності незалежних випробувань

Нехай є два простри подій Ω_1, Ω_2 ,

$$x \in \Omega_1, \quad y \in \Omega_2, \quad (x, y) \in \Omega_1 \times \Omega_2, \quad (y, x) \in \Omega_2 \times \Omega_1,$$

і

$$\Pr(x) = p_1, \quad \Pr(y) = p_2.$$

Якщо події x, y є незалежними в інтуїтивному розумінні, то потрібно подіям (x, y) та (y, x) приписувати ймовірності

$$\Pr(x, y) = \Pr(y, x) = p_1 \cdot p_2.$$

Ще раз підкреслимо, що є математичний термін “незалежні події” і є інтуїтивна уява про незалежні події. Інтуїтивну уяву використовуємо при первісному приписування ймовірності послідовності двох подій. Інтуїтивну уяву можна підсилити дослідями чи якимись міркуваннями. Але дати строге математичне доведення інтуїтивної уяви неможливо. Математичну властивість двох подій “бути незалежними” можна довести. Для цього досить обчислити відповідні ймовірності.

Коли розглядали послідовності випробувань Бернуллі казали, що ймовірність певної послідовності дорівнює

$p^{n_1} \cdot q^{n_2}$, де n_1 — кількість успіхів, а n_2 кількість невдач в цій послідовності, то ми користувалися припущенням, що випробування Бернуллі незалежні.

Підкидаючи монету пишемо 1 або 0 в залежності від того, впаде монета цифрою чи гербом вверх. Запис 0 та 1 є випадковими подіями, що мають ймовірності $\frac{1}{2}$. Вважаємо, що запис наступного числа в послідовності є незалежним від запису попереднього. Тому

$$\Pr(0, 0) = \Pr(0, 1) = \Pr(1, 0) = \Pr(1, 1) = \frac{1}{4}.$$

Подібним чином, вважаючи, що підкидання відбуваються незалежно одне від іншого, одержуємо, що результат трьох підкидань має розподіл ймовірностей

$$\begin{aligned} \Pr(0, 0, 0) &= \Pr(0, 0, 1) = \Pr(0, 1, 0) = \Pr(0, 1, 1) = \\ &= \Pr(1, 0, 0) = \Pr(1, 0, 1) = \Pr(1, 1, 0) = \Pr(1, 1, 1) = \frac{1}{8}, \end{aligned}$$

і, відповідно, результатом n підкидань буде послідовність із нулів та одиниць довжини n . Ймовірність кожної такої елементарної події $a_1, a_2, \dots, a_n$ $1 \leq i \leq n$, $i \in \{0, 1\}$ дорівнює $\frac{1}{2^n}$.

Приклад 3.4 Припустимо, що монета підкидається до тих пір, поки не впаде цифрою вверх. Поставимо питання: яка ймовірність того, що перша одиниця впаде при парному підкиданні.

Простором подій $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$ тут є скінченні послідовності нулів та одиниць, в яких всі елементи, крім останнього, дорівнюють нулю, а останній дорівнює 1.

$$\omega_1 = (1), \omega_2 = (01), \omega_3 = (001), \omega_4 = (0001), \dots$$

Згідно з попередніми міркуваннями, вважаємо, що ймовірність появи послідовності довжини n дорівнює $\frac{1}{2^n}$. Оскільки

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

то розподіл ймовірності у просторі Ω заданий коректно — нескінченною послідовністю підкидань монети бути не може.

Бажана подія та, в якій число n парне. Елементарними подіями є одержання тих послідовностей, в яких n парне. Отже бажана подія $\Omega_1 \subseteq \Omega$ складається із нескінченної кількості елементарних подій. За правилом підрахунку ймовірності події маємо

$$\Pr(\Omega_1) = \sum_{k=1}^{\infty} \frac{1}{2^{2k}}.$$

Використовуючи формулу для знаходження суми членів нескінченної спадної геометричної послідовності маємо

$$\Pr(\Omega_1) = \frac{\frac{1}{4}}{1 - \frac{1}{2^2}} = \frac{1}{4} \cdot \frac{4}{3} = \frac{1}{3}.$$

3.3 Означення умовної ймовірності.

Нехай елементарними подіями в просторі подій Ω_1 є вибір цілого числа від 1 до 100 включно. Тоді вибір парного числа від 1 до 100 є подією $\Omega_2 \subseteq \Omega_1$, а вибір числа від 1 до 100, яке ділиться на 4, це також подія Ω_3 , але вона міститься в Ω_2 (див. рис. 3). Розглядаємо випадок, коли $\Pr(\Omega_2)$ — ймовірність події Ω_2 ненульова, тобто в ній містяться елементарні події з ненульовою ймовірністю. Тоді Ω_2 можна розглядати як новий ймовірнісний

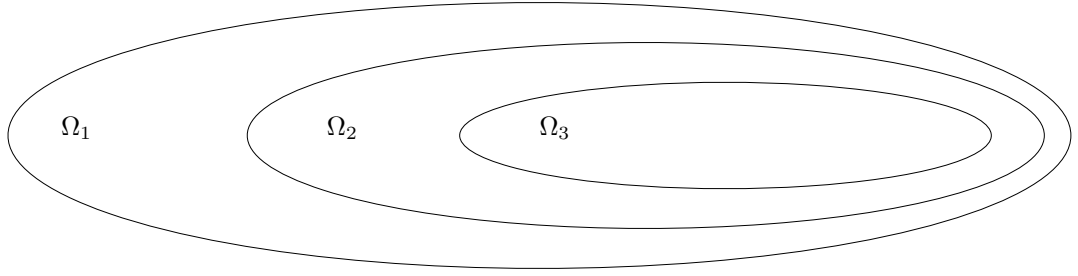


Рис. 3: Ω_3 можна розглядати як подію в просторі подій Ω_1 і в просторі подій Ω_2 .

простір. Щоб природно ввести ймовірність елементарних подій із Ω_2 , потрібно ймовірності всіх елементарних подій із Ω_2 домножити на масштабуючий множник x так, щоб сума ймовірностей всіх елементарних подій із Ω_2 дорівнювала 1. Оскільки ми маємо

$$\sum_{\omega \in \Omega_2} x \cdot \Pr(\omega) = x \sum_{\omega \in \Omega_2} \Pr(\omega) = x \cdot \Pr(\Omega_2) = 1,$$

то масштабуючий множник повинен дорівнювати

$$x = \frac{1}{\Pr(\Omega_2)}.$$

Таким чином, кожна елементарна подія $\omega \in \Omega_3$ одержує дві ймовірності — ймовірність $\Pr(\omega)$ в просторі Ω_1 і $\frac{\Pr(\omega)}{\Pr(\Omega_2)}$ в просторі Ω_2 . Подія Ω_3 в просторі Ω_2 має ймовірність $\frac{\Pr(\Omega_3)}{\Pr(\Omega_2)}$.

Коли Ω_3 розглядається як подія в просторі подій Ω_2 то кажуть що відбувається подія Ω_3 за умови Ω_2 . Ймовірність подій Ω_3 за умови, що відбулася подія Ω_2 , позначається $\Pr(\Omega_3|\Omega_2)$.

Отже, за означенням, коли $\Omega_3 \subseteq \Omega_2$ і $\Pr(\Omega_2) \neq 0$, тоді

$$\Pr(\Omega_3|\Omega_2) = \frac{\Pr(\Omega_3)}{\Pr(\Omega_2)}.$$

Для будь-яких подій $\Omega_2, \Omega_2 \subseteq \Omega_1$ в просторі подій Ω_1 вважається (за означенням)

$$\Pr(\Omega_3|\Omega_2) = \Pr(\Omega_3 \cap \Omega_2|\Omega_2).$$

Використовуючи позначення $A \cap B = AB$ для будь-яких подій A, B в певному ймовірнісному просторі можна писати

$$\Pr(B|A) = \Pr(AB|A) = \frac{\Pr(AB)}{\Pr A}.$$

Визначення 3.2 *Ймовірність*

$$\frac{\Pr(\Omega_2\Omega_3)}{\Pr(\Omega_2)}$$

називають умовною ймовірністю, ймовірністю події Ω_3 за умови події Ω_2 , і позначають $\Pr(\Omega_3|\Omega_2)$.

3.4 Формула повної ймовірності

Теорема 3.1 (Формула повної ймовірності) *Нехай у певному ймовірнісному просторі Ω є подія $A \subseteq \Omega$ і є повна група несумісних подій $B_1, B_2, \dots, B_n$, $n \geq 1$.*

Тоді

$$\Pr(A) = \sum_{i=1}^n \Pr(B_i) \cdot \Pr(A|B_i)). \quad (26)$$

Доведення. Те, що події $B_1, B_2, \dots, B_n$, $n \geq 1$ несумісні і утворюють повну групу, означає

$$\cup_{i=1}^n B_i = \Omega,$$

і

$$i \neq j \Rightarrow B_i \cap B_j = \emptyset, \quad (1 \leq i, j \leq n).$$

Тому

$$A \cap (\cup_{i=1}^n B_i) = A \cap \Omega = A,$$

$$A \cap (\cup_{i=1}^n B_i) = \cup_{i=1}^n (A \cap B_i) = A, \quad (27)$$

Оскільки події $B_1, \dots, B_n$ несумісні, то несумісними будуть також події $A \cap B_1, \dots, A \cap B_n$, що дозволяє застосувати до рівності (27) теорему суми ймовірностей несумісних подій і одержати

$$\Pr(A) = \Pr(\cup_{i=1}^n (A \cap B_i)) = \sum_{i=1}^n \Pr(AB_i). \quad (28)$$

За означенням умовної ймовірності

$$\Pr(A|B_i) = \frac{\Pr(AB_i)}{\Pr(B_i)}$$

і, відповідно, замість $\Pr(AB_i)$ в (28) можна підставити вираз $\Pr(A|B_i) \cdot \Pr(B_i)$ і одержати потрібне — (26)

■

Приклад 3.5 *Нехай є певна кількість червоних та чорних куль, що розташовані в урнах з номерами 1, 2, 3. Відповідно, є чотири події*

A — довільно вибрана куля є червоною;

B_i — довільно вибрана куля лежить в урні з номером $i, i = 1, 2, 3$.

Припустимо, що ймовірність вибору червоної кулі в першій урні дорівнює 0,2, в другій урні дорівнює 0,9, і в третій урні дорівнює 0,5. Припустимо також, що події B_1, B_2, B_3 рівноймовірні.

Чому дорівнює ймовірність події A ?

Оскільки B_1, B_2, B_3 утворюють повну групу несумісних подій, то

$$\Pr(B_1) + \Pr(B_2) + \Pr(B_3) = 1,$$

і, відповідно, $\Pr(B_i) = \frac{1}{3}$ ($i = 1, 2, 3$).

За формулою повної ймовірності маємо

$$\Pr(A) = 0,2 \cdot \frac{1}{3} + 0,9 \cdot \frac{1}{3} + 0,5 \cdot \frac{1}{3} = \frac{8}{15}.$$

Приклад 3.6 Припустимо, що є дві партії деталей, серед яких є браковані. Кількість деталей в обох партіях одна і та ж. Ймовірність того, що деталь із першої партії стандартна (не бракована) дорівнює 0,8, а ймовірність того, що деталь із другої партії виявиться стандартною, дорівнює 0,9.

Вибирається одна деталь із цих двох партій. Яка ймовірність того, що вона виявиться стандартною?

Маємо три події

A — довільно вибрана деталь є стандартною;

B_1 — деталь належить першій партії, $\Pr(B_1) = \frac{1}{2}$;

B_2 — деталь належить другій партії, $\Pr(B_2) = \frac{1}{2}$.

Потрібно знайти ймовірність події A .

За умовою

$$\Pr(A|B_1) = 0,8, \quad \Pr(A|B_2) = 0,9.$$

Формула повної ймовірності дозволяє записати

$$\Pr(A) = \Pr(A|B_1) \cdot \Pr(B_1) + \Pr(A|B_2) \cdot \Pr(B_2) = 0,4 + 0,45 = 0,85.$$

Приклад 3.7 Є дві коробки з дисками. В першій 5 дисків, із них 2 уже використовувалися, а в другій 3, з яких використовувався один. Із другої коробки взяли диск і переклали в першу. Після цього вибирається

один диск із першої коробки. Знайти ймовірність того, що цей вибраний диск із першої коробки не використовувався.

Елементарною подією є вибір диска із першої коробки після того, як вибрали певний диск із другої коробки і перенесли його в першу. Генеральна сукупність Ω складається із 18 елементарних подій — із другої коробки диск можна взяти 3-ма способами, а після цього в першій коробці диск можна взяти 6-ма способами. Оскільки нічого іншого не сказано, то розподіл ймовірності вважається рівномірним, і кожна елементарна подія $\omega \in \Omega$ має ймовірність

$$\Pr(\omega) = \frac{1}{18}.$$

Сприятливих елементарних подій 11: по 4 при перенесенні невикористаного диска і 3 при перенесенні використаного диска. Отже шуканою ймовірністю буде $\frac{11}{18}$.

Розв'яжемо цю ж задачу з використанням формули повної ймовірності. Для цього поозначимо події

B_1 — “із другої коробки вибрано невикористаний диск”;

B_2 — “із другої коробки вибраний використаний диск”;

A — “Після перенесення із першої коробки вибраний невикористаний диск”.

Маємо

$$\Pr(B_1) = \frac{2}{3}, \quad \Pr(B_2) = \frac{1}{3},$$

$$\Pr(A|B_1) = \frac{4}{6} = \frac{2}{3}, \quad \Pr(A|B_2) = \frac{3}{6} = \frac{1}{2},$$

За формулою повної ймовірності одержуємо відповідь

$$\Pr(A) = \Pr(B_1) \cdot \Pr(A|B_1) + \Pr(B_2) \cdot \Pr(A|B_2) = \frac{2}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{11}{18}.$$

3.5 Формули Баєса

Нехай Ω ймовірнісний простір, і в ньому є дві події $A, B \subseteq \Omega$. За означенням умовної ймовірності

$$\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)}, \quad \Pr(B|A) = \frac{\Pr(AB)}{\Pr(A)}.$$

Звідси маємо

$$\Pr(AB) = \Pr(A|B) \cdot \Pr(B) = \Pr(B|A) \cdot \Pr(A).$$

Розділивши останню рівність на $\Pr(A)$ одержуємо рівність

$$\Pr(B|A) = \frac{\Pr(A|B) \cdot \Pr(B)}{\Pr(A)}. \quad (29)$$

Доведена формула (29) називається формулою Байєса для двох подій.

Нехай Ω ймовірнісний простір, $B_1, B_2, \dots, B_n, n \geq 1$ повна група незалежних подій в цьому просторі і $A \subseteq \Omega$ — подія. Для кожного $i = 1, 2, \dots, n$ перепишемо формулу (29), замінивши B на B_i і скористаємося формулою повної ймовірності (26). Одержимо

$$\Pr(B_i|A) = \frac{\Pr(A|B_i) \cdot \Pr(B_i)}{\sum_{k=1}^n (\Pr(B_k) \cdot \Pr(A|B_k))}. \quad (30)$$

Одержані формули (30) називають формулами Байєса для повного набору несумісних подій.

Томас Баєс (Thomas Bayes) народився в 1702 в Лондоні, з 1720 року працював священиком приходу в містечку Танбрідж Уелс — біля 50 км від Лондона. За

життя анонімно опублікував одну математичну працю в захист Ньютонівського диференційного числення, а під своїм іменем не надрукував жодної наукової праці. В 1742 році був обраний членом Лондонського Королівського товариства. Після смерті Томаса Баєса його товариш Річард Прайс помітив серед паперів фундаментальну працю по теорії ймовірності “Есе про розв’язування проблем в теорії випадкових подій”, направив цю роботу в Лондонське Королівське Товариство, і в 1764 році “Есе про розв’язування проблем в теорії випадкових подій” було опубліковане в працях цього товариства. З цього часу починається всесвітня слава Томаса Байєса. Помер Томас Байєс 17 квітня 1761 року.⁵

Приклад 3.8 *При грі в покер гравцям роздають карти і ті, оцінюючи одержані карти, приймають рішення — продовжити гру чи припинити. Ймовірність того, що даний гравець в конкретній роздачі продовжить гру називається впін’ом⁶*

Розглянемо наступну ситуацію. Статистика на сайті в інтернеті, де грають в покер, показує, що 10 відсотків гравців мають більший впін ніж 0.52 (позначимо цю групу гравців L , і, відповідно, 90 відсотків мають впін менше 0.52. Позначимо другу групу гравців через M . Статистика показує, що при розрахунках можна прийняти впін гравця із L ріним 0.62, а впін гравця із M — 0,26 .

На сайт заходить гравець і в перших 6 задачах карт 5 разів приймає рішення продовжити гру. Яка ймовірність того, що цей гравець належить групі L ?

Для розбору ситуації позначимо:

⁵Інформація взята із <http://old.computerra.ru/online/jack/12150/>

⁶утворено першими буквами Voluntary put \$ in pot — добровільно поклади долар в горщик

B_1 — гравець належить групі L ;
 B_2 — гравець належить групі M ;
 A — гравець в 6-и роздачах карт 5 разів приймає рішення продовжити гру.
 За умовою,

$$\Pr(B_1) = 0,1, \quad \Pr(B_2) = 0,9.$$

Ймовірності $\Pr(B_1), \Pr(B_2)$ — це апіорні⁷ оцінки припущень (гіпотез) — “гравець, що прийшов, належить десятивідсотковій групі L ” та “гравець, що прийшов, належить дев’яностовідсотковій групі M ”.

За умовою, якщо гравець належить групі L , то ймовірність елементарної події — “він продовжив гру після роздачі” дорівнює 0,62, а ймовірність елементарної події “він припинив гру після роздачі” дорівнює $1 - 0,62 = 0,38$. Вважаємо кожне наступне продовження чи не продовження гри після роздачі незалежним від попереднього. Тому конкретна послідовність із 6 роздач, в яких гравець 5 разів продовжив гру і один раз не продовжив гру, має ймовірність

$$\Pr(A|B_1) = (0.62)^5 \cdot 0.38 = 0.03481304762.$$

Подібним чином підраховуємо ймовірність $\Pr(A|B_2)$:

$$\Pr(A|B_2) = (0.26)^5 \cdot 0.74 = 0.000879221824.$$

Оскільки $B_1 B_2$ повний набір незалежних подій, то ми можемо скористатися формою повної ймовірності для знаходження ймовірності $\Pr(A)$:

$$\begin{aligned} \Pr(A) &= \Pr(A|B_1)\Pr(B_1) + \Pr(A|B_2)\Pr(B_2) = \\ &= 0.03481304762 \cdot 0.1 + 0.000879221824 \cdot 0.9 = 0.004272604404. \end{aligned}$$

⁷а priori — незалежно від досвіду, в даному випадку, незалежно від знань, одержаних після 6 роздач карт

Для одержання потрібних нам умовних ймовірностей виписуємо необхідні формули Баєса

$$\Pr(B_1|A) = \frac{\Pr(A|B_1) \cdot \Pr(B_1)}{\Pr(A)}, \quad \Pr(B_2|A) = \frac{\Pr(A|B_2) \cdot \Pr(B_2)}{\Pr(A)},$$

і підставляємо в них потрібні значення

$$\Pr(B_1|A) = \frac{(0.03481304762) \cdot 0.1}{0.004272604404} = 0.8147968857,$$

і, відповідно,

$$\Pr(B_2|A) = \frac{\Pr(A|B_2) \cdot \Pr(B_2)}{\Pr(A)} = \frac{0.000879221824 \cdot 0.9}{0.004272604404} = 0.1852031143.$$

Ймовірності $\Pr(B_1|A)$, $\Pr(B_2|A)$ називають апостеріорними⁸. Крім того, події B_1 , B_2 в такому мовному оточенні називають гіпотезами.

Отже одержаний результат можна переформулювати так. Щодо гравця, що прийшов на сайт, є дві опріорні гіпотези — або він належить групі L (ймовірність цієї гіпотези 0.1), або він належить групі M (ймовірність цієї гіпотези 0.9). Після 6 роздач відбувся перерахунок ймовірностей, — нові ймовірності ґрунтуються на подіях, що відбулися, на досвіді, це апостеріорні ймовірності. І в нашому випадку, ймовірність того, що гравець належить групі L уже становить біля 81 відсотка, а ймовірність того, що він належить групі M , становить лише біля 19 відсотків.

Приклад 3.9 *Із 20 студентів, що прийшли на іспит, 8 підготовані відмінно, 6 - добре, 4 - задовільно і 2 -*

⁸a posteriori — що ґрунтується на досвіді

погано, незадовільно. На іспиті студентам пропонується 40 екзаменаційних білетів. Відмінно підготований студент знає відповіді на будь-який білет. Відповідно, добре підготований знає відповіді до 30 білетів, задовільно — до 20 білетів, а незадовільно підготований знає відповіді лише до 10 білетів.

Студент взяв білет, на який він знав відповіді. Яка ймовірність того, що цей студент підготований незадовільно.

Ввадимо 4 гіпотези

B_1 — студент, що витяг білет, підготований відмінно;

B_2 — студент, що витяг білет, підготований добре;

B_3 — студент, що витяг білет, підготований задовільно;

B_4 — студент, що витяг білет, підготований незадовільно.

За умовою маємо чотири апріорні ймовірності цих гіпотез

$$\Pr(B_1) = \frac{8}{20} = 0.4, \quad \Pr(B_2) = \frac{6}{20} = 0.3, \quad \Pr(B_3) = \frac{4}{20} = 0.2, \quad \Pr(B_4) = \frac{2}{20} = 0.1$$

Позначимо через

A — подію “студент, що взяв білет, знає відповіді на всі питання цього білету.”

За умовою, ми знаємо 4 умовні ймовірності:

$$\begin{aligned} \Pr(A|B_1) &= \frac{40}{40} = 1, & \Pr(A|B_2) &= \frac{30}{40} = 0.75, \\ \Pr(A|B_3) &= \frac{20}{40} = 0.5, & \Pr(A|B_4) &= \frac{10}{40} = 0.25, \end{aligned}$$

За формулою повної ймовірності ми можемо знайти ймовірність події A :

$$\begin{aligned}\Pr(A) &= \Pr(A|B_1) \cdot \Pr(B_1) + \Pr(A|B_2) \cdot \Pr(B_2) + \Pr(A|B_3) \cdot \Pr(B_3) + \Pr(A|B_4) \cdot \Pr(B_4) \\ &= 0.4 + 0.75 \cdot 0.3 + 0.5 \cdot 0.2 + 0.25 \cdot 0.1 = 0.625 + 0.125 = 0.75\end{aligned}$$

Одержані дані дозволяють обчислити за допомогою формул Баєса апостеріорні ймовірності гіпотез

$$\Pr(B_1|A) = \frac{\Pr(A|B_1) \cdot \Pr(B_1)}{\Pr(A)} = \frac{1 \cdot 0.4}{0.75} = 0.53333,$$

$$\Pr(B_2|A) = \frac{\Pr(A|B_2) \cdot \Pr(B_2)}{\Pr(A)} = \frac{0.75 \cdot 0.3}{0.75} = 0.3,$$

$$\Pr(B_3|A) = \frac{\Pr(A|B_3) \cdot \Pr(B_3)}{\Pr(A)} = \frac{0.5 \cdot 0.2}{0.75} = 0.133333,$$

$$\Pr(B_4|A) = \frac{\Pr(A|B_4) \cdot \Pr(B_4)}{\Pr(A)} = \frac{0.25 \cdot 0.1}{0.75} = 0.033333,$$

Приклад 3.10 Обчислювальній системі потрібно зрозуміти, що означає слово “коса”. Вона має три гіпотези — це або спосіб оформлення жіночого волосся, або інструмент для зрізування трави, або форма берега. Ці три розуміння мають певні ймовірності появи в тексті — апріорні оцінки. Але подія “в контексті зустрілися слова *заплітати*, *розчісувати*,” може змінити ці апріорні ймовірності.

Сучасні розробники програмного забезпечення часто використовують теорему Баєса — Баєсівський статистичний аналіз.

4 Випадкова величина, ймовірність випадкової величини

4.1 Випадкова величина

Функція

$$f : \Omega \rightarrow \mathbb{R},$$

що визначена на просторі елементарних подій Ω , називається випадковою величиною.

В подальшому ми будемо мати справу тільки з випадковими величинами і для скорочення називатимемо їх просто величинами.

Приклад 4.1 Візьмемо простором елементарних подій розташування n однакових куль по m перенумерованих урнах. Кожній події можна поставити у відповідність число — кількість куль, що попали в першу урну. Це буде приклад величини.

Приклад 4.2 Підкинута вгору куля падає на землю — це елементарна випадкова подія, а відстань від місця підкидання до місця падіння буде випадковою величиною.

Приклад 4.3 Є дві партії деталей. Вибираємо одну деталь — це елементарна випадкова подія. А номер партії, якій належить взята деталь — це величина.

Нехай маємо простір випадкових подій Ω і $f : \Omega \rightarrow \mathbb{R}$ — випадкова величина. Тоді для кожного $a \in \mathbb{R}$ можна розглядати подію $\Omega_a \subseteq \Omega$, що визначається умовою

$$\omega \in \Omega_a \Leftrightarrow f(\omega) = a. \quad (31)$$

Приклад 4.4 Нехай ймовірнісний простір складається із елементарних подій — вибору однієї книги із 20. Ко-

жній книзі ставимо у відповідність її висоту в сантиметрах. Отже висота книги — це величина у просторі книг. Значення 21 цієї величини задає подію — вибір книги висотою 21 см.

4.2 Ймовірність випадкової величини.

Визначення 4.1 Ймовірністю $Pr(a)$ заданого значення a величини f називають ймовірність $Pr(\Omega_a)$ відповідної події

$$\Omega_a = \{\omega \in \Omega | f(\omega) = a\}.$$

Далі будемо працювати і з випадками, коли простором елементарних випадкових подій є можливі значення певної величини. Наприклад, при підкиданні кубика нас частіше буде цікавити номер, що написаний на верхній грані, а не колір цієї грані.

Величина $f : \Omega \rightarrow \mathbb{R}$ задає на Ω відношення еквівалентності $\sim$:

$$\omega_1 \sim \omega_2 \Leftrightarrow f(\omega_1) = f(\omega_2). \quad (32)$$

Відношення еквівалентності (32) задає розбиття ймовірнісного простору, тобто задає події $\Omega_a \subseteq \Omega$, що мають дві визначальні властивості

$$\cup_{a \in \Omega} \Omega_a = \Omega; \quad (33)$$

$$\Omega_a \cap \Omega_b \neq \emptyset \Rightarrow \Omega_a = \Omega_b. \quad (34)$$

Із властивостей (33), (34) випливає, що

$$\sum_{a \in f(\Omega)} Pr(a) = 1$$

і значення величини можна вважати простором елементарних випадкових подій, простором елементарних подій.

Приклад 4.5 Якщо елементарними подіями є послідовності випробувань Бернуллі, то їм можна ставити у відповідність довжину (якщо вони скінченні), або кількість Успіхів в послідовності, або кількість Невдач, що передують Успіхові, або номер місця, на якому зустрівся перший Успіх та подібне. Все це — випадкові величини.

Приклад 4.6 Якщо подіями є вибір однієї карти з колоди гральних карт, то випадковою величиною може бути номер за зростом. Тоді

$$Pr(4) = Pr(\{5\spadesuit, 5\clubsuit, 5\diamondsuit, 5\heartsuit\}) = \frac{1}{13}.$$

Випадковою величиною може бути також номер N масті вибраної карти. Тоді $N(\spadesuit) = 1$, $N(\clubsuit) = 2$, $N(\diamondsuit) = 3$, $N(\heartsuit) = 4$, і

$$Pr(4) = Pr(\{2\heartsuit, 3\heartsuit, 4\heartsuit, 5\heartsuit, 6\heartsuit, 7\heartsuit, 8\heartsuit, 9\heartsuit, 10\heartsuit, B\heartsuit, D\heartsuit, K\heartsuit, T\heartsuit\}) = \frac{1}{4}.$$

Приклад 4.7 У випробуваннях Бернуллі зазвичай значення величини розуміються як величини виграшу (чи програшу — від'ємного виграшу), яким приписуються ймовірності p, q . Припустимо, що ви є учасником лотереї, в якій вартість одного лотерейного білета 1 гривня, величина виграшу за цим білетом складає 100 гривень, а ймовірність p виграшу за цим білетом складає 0.009. Отже тут виграш — величина Успіху $f(\text{Успіх}) = 100$, програш — величина Невдачі $f(\text{Невдача}) = -1$, ймовірність Невдачі дорівнює 0.991:

$$Pr(100) = 0.009, \quad Pr(-1) = 0.991.$$

Приклад 4.8 При підкиданні гральної кістки грані нумерують довільним чином числами $1, 2, 3, 4, 5, 6$, і величиною грані є її номер. В припущенні, що центр ваги в кістці не збитий, для цієї величини беруть рівномірний розподіл $Pr(i) = \frac{1}{6}$, $i = 1, 2, 3, 4, 5, 6$.

Приклад 4.9 Функцію розподілу⁹ ймовірності випадкової величини зручно зображати на координатній площині. Наведемо приклад біномного розподілу величини $n = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$, $Pr(n) = C_{10}^n p^n q^{10-n}$, $p = \frac{1}{3}$, $q = \frac{2}{3}$:

$$\begin{aligned} Pr(0) &= 0.01734152992, Pr(1) = 0.08670764958, Pr(2) = 0.1950922116, \\ Pr(3) &= 0.2601229487, Pr(4) = 0.2276075801, Pr(5) = 0.1365645481, \\ Pr(6) &= 0.05690189504, Pr(7) = 0.01625768430, Pr(8) = 0.003048315806, \\ Pr(9) &= 0.0003387017562, Pr(10) = 0.00001693508779. \end{aligned}$$

Графік цього розподілу наведений на рис. 4.2.

Нехай на просторі випадкових подій Ω задана величина X , тобто задане відображення $X : \Omega \rightarrow \mathbb{R}$. Візьмемо певне значення $a \in \mathbb{R}$. Теретико-ймовірнісна традиція рекомендує подію “величина X ставить у відповідність елементарної поді $\omega \in \Omega$ значення a ” позначати

$$X = a.$$

В сучасному теретико-множинному записі $X = a$ означає

$$\{\omega \in \Omega | X(\omega) = a.\}$$

Запис $Pr(X = a, Y = b)$ для величин X, Y означає, що є дві величини X, Y на одному ймовірнісному просторі Ω і розглядається ймовірність події $\Omega_1 \subseteq \Omega$:

⁹Нагадаємо, що в загальних розділах теорії ймовірності під функцією розподілу розуміють дещо інше. Про це буде йти розмова в наступному розділі “кумулятивна функція розподілу”.

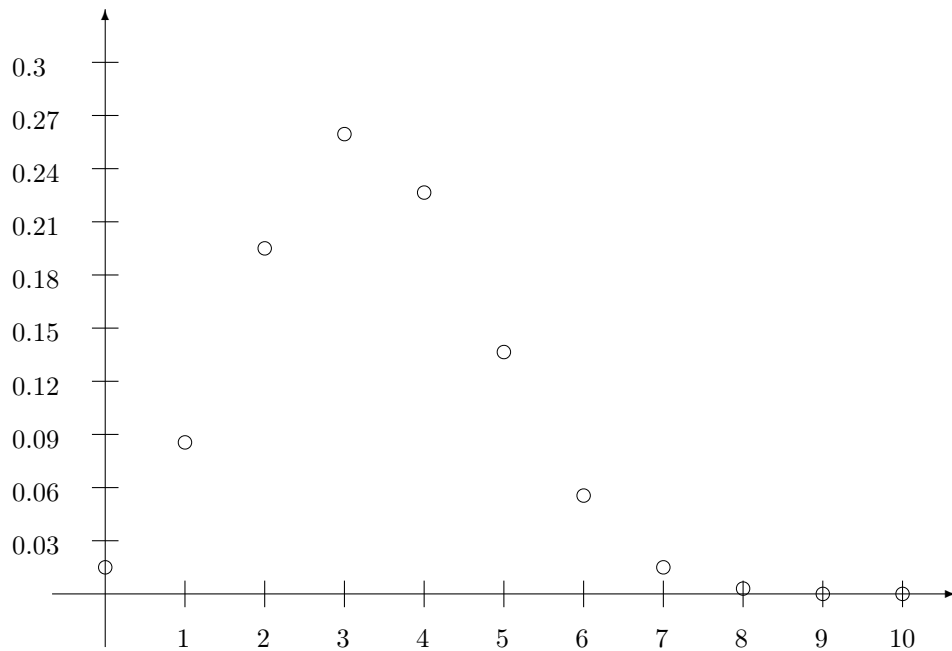


Рис. 4: Біномний розподіл величини $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

$$\Omega_1 = \{\omega \in \Omega | X(\omega) = a\} \cap \{\omega \in \Omega | Y(\omega) = b\}.$$

4.3 Інтегральна (кумулятивна) функція розподілу ймовірності

Крім введеної функції розподілу ймовірності (у просторі подій) розглядають також ще одну функцію розподілу ймовірності (випадкової величини).

Підхід, коли функція $f : \Omega \rightarrow [0, 1]$ розподілу ймовірності на просторі випадкових подій означається як на стор. 7, в багатьох випадках нелегко у використанні.

Щоб побачити це, розглянемо приклад.

Приклад 4.10 Гра “гра в рулетку” полягає в тому, що в круг, що розділений на n однакових секторів, вкидається кулька. Елементарною подією є зупинка

кульки в певному секторі. Ймовірність кожної елементарної події дорівнює $\frac{1}{n}$ (рівномірний розподіл). Виграє той сектор, в якому зупинилася рулетка.

Подвоюючи кількість секторів при грі в рулетку ми вдвічі зменшуємо ймовірність елементарної події. Продовжуючи подвоєння секторів ми прийдемо до граничного випадку — сектори стають радіусами, а ймовірність елементарної події стає нулем. Але тоді ймовірність події перестає дорівнювати сумі ймовірностей елементарних подій, що входять до неї.

Наведений приклад обґрунтовує тезу про незручність використання функції розподілу, що введена на стор. 7) за межами дискретної теорії ймовірності. В роботі із випадковими величинами значні зручності надає так звана кумулятивна або інтегральна функція $F : \mathbb{R} \rightarrow [0, 1]$ розподілу. Вона визначаються за допомогою правила:

$$F(a) = \Pr(x \leq a). \quad (35)$$

Кумулятивну функцію розподілу ймовірності неможливо використовувати тоді, коли на просторі подій немає лінійного порядку. Графік кумулятивної функції біномного розподілу ймовірності випадкової величини з параметрами $n = 10, p = \frac{1}{3}$ зображений на рис. 4.3. А графік функції розподілу цієї величини зображений на рис. 4.2.

Далі кумулятивна функція розподілу не буде використовуватися.

4.4 Випадок двох величин на одному ймовірнісному просторі

На одному і тому ж ймовірнісному просторі можуть бути задані дві і більше величин. Приклад 4.6 показує це.

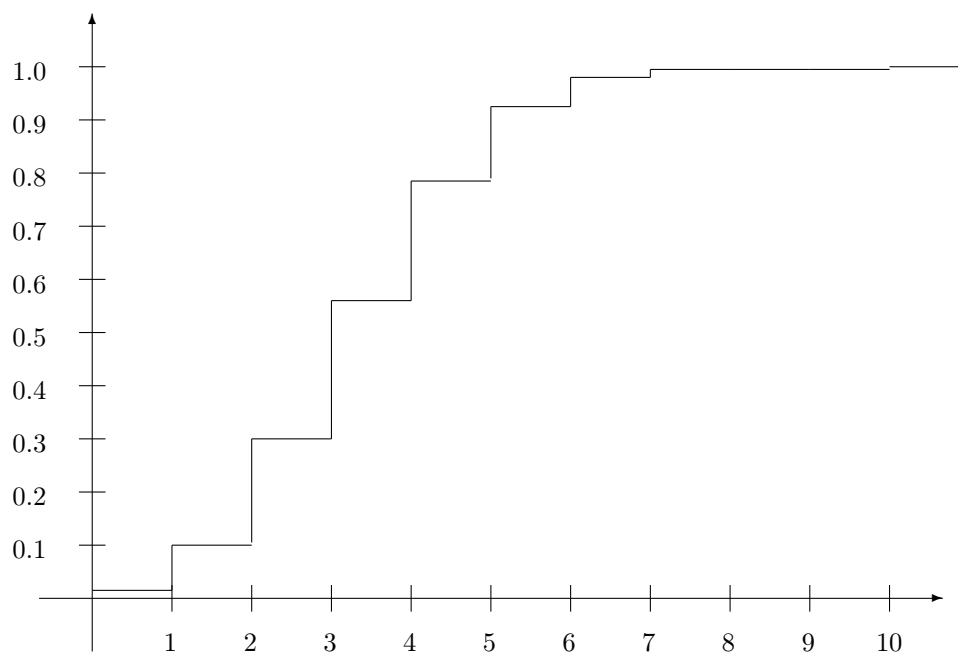


Рис. 5: Інтегральна (кумулятивна) функція розподілу — біномний розподіл величини $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

Можна також із уже відомих величин будувати нові. Так, коли на ймовірнісному просторі Ω задані величини $f, g : \Omega \rightarrow \mathbb{R}$, то можна побудувати нові величини h , наприклад

$$h = 2f, -f, f^2, f + g, f \cdot g, 2f + 3g \quad (36)$$

та подібне. Значення функцій h , наведених в (36), від елементарної події ω обчислюється за правилом

$$\begin{aligned} (2f)(\omega) &= 2 \cdot (f(\omega)), & (-f)(\omega) &= -(f(\omega)), \\ (f^2)(\omega) &= (f(\omega))^2, & (f + g)(\omega) &= (f(\omega) + g(\omega)), \\ (fg)(\omega) &= f(\omega) \cdot g(\omega), & (2f + 3g)(\omega) &= 2 \cdot f(\omega) + 3 \cdot g(\omega). \end{aligned}$$

Приклад 4.11 *Кидають дві гральні кістки — скажемо, зелену та жовту. Випавший парі граней ставимо у відповідність те число, яке випало на зеленій кістці — це одна величина. Але цьому ж випадінню можна*

поставити у відповідність число, яке випало на жовтій грані — це друга величина. Проте випадінню двох чисел природно поставити у відповідність суму очок, що випали на кістках, — це сума двох величин.

Дві величини X, Y на ймовірнісному просторі Ω називаються незалежними, якщо для будь-яких $a, b \in \mathbb{R}$ події $X = a$ та $Y = b$ є незалежними, тобто

$$\Pr(X = a, Y = b) = \Pr(X = a) \cdot \Pr(Y = b).$$

Теорема 4.1 *Нехай X, Y — дві незалежні величини на ймовірнісному просторі Ω , A — множина значень величини X , які мають ненульові ймовірності, і B — множина значень величини Y , що мають ненульові ймовірності. Тоді множиною C значень величини XY , які мають ненульові ймовірності, буде*

$$C = \{a \cdot b | a \in A, b \in B\}. \quad (37)$$

Доведення. Виберемо довільний елемент $c \in C$. За означенням $\Pr(XY = c) \neq 0$, і для деякого $\omega_0 \in \Omega$

$$(XY)(\omega_0) = c, \quad X(\omega_0) \cdot Y(\omega_0) = c, \quad \Pr(\omega_0) > 0.$$

Позначимо

$$X(\omega_0) = a, \quad Y(\omega_0) = b.$$

В цих позначеннях

$$\Pr(X = a) \geq \Pr(\omega_0) > 0, \quad \Pr(Y = b) \geq \Pr(\omega_0) > 0, \quad a \cdot b = c,$$

можна стверджувати, що $c = a \cdot b$ для деяких $a \in A, b \in B$, і

$$C \subseteq \{a \cdot b | a \in A, b \in B\}. \quad (38)$$

При цьому незалежність випадкових величин X, Y не використовувалась.

Нехай $a \in A, b \in B$. Тоді за означенням множин A, B буде $\Pr(X = a) \neq 0, \Pr(Y = b) \neq 0$. За означенням незалежних випадкових величин буде

$$\Pr(X = a, Y = b) = \Pr(X = a) \cdot \Pr(Y = b) \neq 0.$$

Звідси випливає, що $\Pr(X = a, Y = b) \neq 0$, і для деякого $\omega_0 \in \Omega$ буде

$$X(\omega) = a, \quad Y(\omega) = b, \quad \Pr(\omega) \neq 0.$$

Звідси випливає, що

$$\Pr(XY = a \cdot b) \geq \Pr(\omega) > 0, ab \in C$$

і

$$C \supseteq \{a \cdot b | a \in A, b \in B\}. \quad (39)$$

Із включень (38), (39) випливає рівність (37). ■

5 Характеристики розподілу ймовірності випадкової величини

5.1 Нескінченні суми дійсних чисел

Пояснимо, що будемо мати на увазі, коли говоритимемо про нескінченну суму дійсних доданків і писатимемо

$$a = \sum_{i \in I} x_i, \quad (40)$$

де I — певна, можливо нескінченна, множина, а $x_i (i \in I)$ є дійсними числами.

Розглядаємо скінченні підмножини $J \subseteq I$ і, відповідно, скінченні суми

$$\sum_{i \in J} x_i.$$

Множину таких скінченних сум позначимо через S . Нехай

$$u = \sup\{s | s \in S\}$$

— точну верхню межу скінченних сум, і

$$v = \inf\{s | s \in S\}$$

— точну нижню межу скінченних сум. Якщо I — скінченна множина, то u є сумою додатних доданків, а v — сума від'ємних доданків. Якщо I є нескінченною множиною, то одне із чисел u, v або обидва можуть не існувати. Якщо обидва числа u, v існують, то вважаємо, що сума (40) існує і

$$a = u + v.$$

Якщо ж бодай одно із чисел u, v не існує, тоді і сума (40) не існує.

Зауважимо, що сума порожньої множини доданків скінченна, сума порожньої множини доданків вважається рівною 0, і тому, коли числа u, v існують, то вони задовольняють нерівності

$$u \geq 0, \quad v \leq 0.$$

Відомо, що коли існують дві суми

$$a = \sum_{i \in I} x_i, \quad b = \sum_{j \in J} y_j,$$

то існує також сума

$$c = \sum_{i \in I, j \in J} x_i \cdot y_j$$

і

$$c = a \cdot b.$$

Згаданими фактами будемо користуватися як відомими.

5.2 Математичне очікування

Припустимо, що ви на постійній основі вкладаєте раз за разом гроші (70 гривень) в певну операцію, з імовірністю 0.15 операція закінчується вдало і ви одержуєте прибуток в 1000 гривень. Якщо операція закінчилася невдало, ваші гроші просто пропадають. Якщо операція проводиться багато разів, то ви маєте право очікувати, що імовірність успіху (операція закінчилася вдало) приблизно дорівнює частоті появи успіхів — якщо операція проводиться n разів, із яких $k \leq n$ — успішні, то

$$\frac{k}{n} \approx 0.15.$$

Таким чином, здійснивши n операцій ви вкладаєте $70n$ гривень і одержуєте прибуток $1000k$. Підсумковий прибуток за n операцій складає

$$(1000 - 70)k - 70(n - k).$$

Розділивши підсумковий прибуток за n операцій на кількість операцій ми одержуємо очікуваний виграш V за одну операцію:

$$V = \frac{(1000 - 70)k - 70(n - k)}{n} = 1000\frac{k}{n} - 70 = 930 \cdot 0.15 - 70 \cdot 0.85 = 80.$$

В наведеному прикладі ми маємо двоелементний ймовірнісний простір $\{Y, H\}$, на якому визначена величина X виграшу

$$X(Y) = 930, \quad X(H) = -70,$$

з функцією розподілу

$$\Pr(930) = 0.15, \quad \Pr(-70) = 0.85.$$

Очікуваний виграш V в одній операції обраховується за формулою

$$V = 930 \cdot \Pr(930) + (-70) \cdot \Pr(-70) = 80.$$

Сказане повинне пояснити, що стоїть за наступним означенням

Визначення 5.1 *Нехай випадкова величина X на ймовірнісному просторі Ω приймає значення $x_1, x_2, \dots$ з відповідною ймовірністю $p_1 = \Pr(X = x_1), p_2 = \Pr(X = x_2), \dots$. Математичним очікуванням¹⁰ величини X називається число*

$$E(X) = x_1 p_1 + x_2 p_2 + \dots = \sum_{i \geq 1} x_i p_i. \quad (41)$$

Якщо сума (41) не існує, то кажуть, що величина X не має математичного очікування

Позначення $E(X)$ йде від першої букви англійського слова expectation, яке означає “очікування”. В радянській і в російськомовній літературі для математичного очікування величини X використовується позначення $M(X)$.

Оскільки

$$\Pr(X = x_i) = \sum_{\omega \in \Omega, X(\omega)=x_i} \Pr(\omega), \quad \sum_{i \geq 1} x_i \Pr(X = x_i) = \sum_{\omega \in \Omega, X(\omega)=x_i} x_i \Pr(\omega),$$

то означення (41) можна переписати у вигляді

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \Pr(\omega). \quad (42)$$

Приклад 5.1 *Нехай величина X має рівномірний розподіл, тобто кожне можливе значення $x_i (i = 1, 2, \dots, n)$*

¹⁰Відношення спільноти до терміну “математичне очікування” неоднозначне. Інтернет-енциклопедія Вікіпедія пропонує єдиний термін — “математичне сподівання”, там немає математичного очікування. В інтернеті започаткована дискусія з пропозицією замінити “сподівання” на “очікування”, оскільки в слові сподівання є додатковий недоречний відтінок мрійливих надій. Там же наведені джерела, де використовуються той чи інший терміни, наведені частотні характеристики їх вживання.

Автор пропонує вважати терміни “математичне очікування” та “математичне сподівання” синонімічними, подібно до пар детермінант — визначник, векторний простір — лінійний простір, і вживати той чи інший термін в залежності від мовного смаку — на розуміння текстів це не вплине.

має ймовірність $Pr(X = x_i) = \frac{1}{n}$. Математичним очікуванням цієї величини буде середнє арифметичне

$$E(X) = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Це видно із означення.

Приклад 5.2 Нехай величина X має біномний розподіл, тобто для кожного $\omega \in \Omega$ $X(\omega) \in \{0, 1, 2, \dots, n\}$, $Pr(X = i) = C_n^i p^i q^{n-i}$, $p, q > 0$, $p + q = 1$. Математичним очікуванням $E(X)$ цієї величини буде np .

Дійсно, за означенням

$$E(X) = \sum_{i=0}^n i \cdot C_n^i p^i q^{n-i}. \quad (43)$$

Зауважимо,

$$i \cdot C_n^i = i \cdot \frac{n \cdot (n-1) \cdot \dots \cdot (n-i+1)}{1 \cdot 2 \cdot \dots \cdot (i-1) \cdot i} = n \frac{(n-1)(n-2) \dots (n-i+1)}{1 \cdot 2 \cdot \dots \cdot (i-1)} = n C_{n-1}^{i-1}.$$

Підставимо одержане значення для $i \cdot C_n^i$ в (43), і одержимо потрібне:

$$E(X) = \sum_{i=1}^n n \cdot C_{n-1}^{i-1} p^i q^{n-i} = np \sum_{k=0}^{n-1} C_{n-1}^k p^k q^{n-1-k} = np(p+q)^{n-1} = np.$$

Приклад 5.3 Нехай величина X має розподіл Пуассона, тобто для кожного $\omega \in \Omega$ $X(\omega) \in \{0, 1, 2, \dots\}$, $Pr(k) = p(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$. Математичним очікуванням $E(X)$ цієї величини буде λ .

Дійсно, за означенням

$$E(X) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda}. \quad (44)$$

Зауважимо,

$$k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda p(k-1, \lambda).$$

Підставимо одержане значення для $k \cdot \frac{\lambda^k}{k!} e^{-\lambda}$ в (44), і одержимо потрібне:

$$E(X) = \lambda \sum_{k=1}^{\infty} p(k-1, \lambda) = \lambda.$$

Теорема 5.1 *Нехай на ймовірнісному просторі задані дві величини X та Y . Тоді*

$$E(X + Y) = E(X) + E(Y). \quad (45)$$

А якщо ці величини незалежні, то

$$E(XY) = E(X) \cdot E(Y). \quad (46)$$

Також для будь-якого числа $c \in \mathbb{R}$

$$E(c \cdot X) = c \cdot E(X). \quad (47)$$

Доведення. Справді,

$$E(X+Y) = \sum_{\omega \in \Omega} (X+Y)(\omega) \Pr(\omega) = \sum_{\omega \in \Omega} (X(\omega)+Y(\omega)) \Pr(\omega) = E(X)+E(Y).$$

Для незалежних величин X, Y позначимо через A множину можливих значень X , а через B — множину можливих значень випадкової величини Y . В цих позначеннях можна записати

$$E(XY) = \sum_{\omega \in \Omega} XY(\omega) \Pr(\omega) = \sum_{a \in A, b \in B} a \cdot b \cdot \Pr(X = a, Y = b) =$$

$$\sum_{a \in A, b \in B} a \cdot b \cdot \Pr(X = a) \cdot \Pr(Y = b) = \sum_{a \in A} a \cdot \Pr(X = a) \cdot \sum_{b \in B} b \cdot \Pr(Y = b) = E(X) \cdot E(Y).$$

Для довільного $c \in \mathbb{R}$

$$E(cX) = \sum_{\omega \in \Omega} cX(\omega) \Pr(\omega) = c \sum_{\omega \in \Omega} X(\omega) \Pr(\omega) = cE(X).$$

■

5.3 Розсіяння (дисперсія)

Латинське слово *dispersio* означає розсіяння. В теорії ймовірності розглядається міра розсіяння випадкової величини навколо свого середнього значення, навколо математичного очікування. За означенням

Дисперсією випадкової величини X , що має математичне очікування $E(X)$, називається математичне очікування величини $(X - E(X))^2$. Дисперсія позначається через $Var(X)$ ¹¹,

Отже, використовуючи символіку, можна записати

$$Var(X) = E((X - E(X))^2).$$

Додатний квадратний корінь із дисперсії називають стандартним (або середнім квадратичним) відхиленням.

Приклад 5.4 *Нехай на ймовірнісному просторі Ω задана стала величина $c \in \mathbb{R}$, тобто*

$$\forall \omega \in \Omega \ c(\omega) = c.$$

Тоді

$$Pr(c(\omega) = c) = \sum_{\omega \in \Omega} Pr(\omega) = 1,$$

$$E(c) = c \cdot Pr(c(\omega) = c) = c \cdot 1 = c,$$

$$Var(c) = E((c - c)^2) = E(0) = 0,$$

і стандартне відхилення дорівнює нулю.

¹¹Від англійського слова *variance*, яке означає “дисперсія”

Теорема 5.2 Якщо випадкова величина X має математичне очікування μ , то дисперсію $Var(X)$ можна обчислювати за формулою

$$Var(X) = E(x^2) - \mu^2.$$

Доведення. Дійсно,

$$Var(X) = E(X - \mu)^2 = E(X^2 - 2\mu X + \mu^2). \quad (48)$$

Використовуючи теорему (5.1), перепишемо праву частину рівності (48) у вигляді

$$E(X - \mu)^2 = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2.$$

Цим завершується доведення теореми. ■

Приклад 5.5 Дисперсія величини з розподілом Пуассона з параметром λ , дорівнює математичному очікуванню і дорівнює λ .

Справді, обчислення дисперсії величини X з розподілом Пуассона $p(k, \lambda)$ показує

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} k \cdot \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \\ &= \lambda \sum_{k=1}^{\infty} (k-1) \cdot \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} + \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda^2 + \lambda. \end{aligned} \quad (49)$$

Тому

$$Var(X) = E(X^2) - E(X)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda.$$

Теорема 5.3 Якщо дві величини на ймовірнісному просторі незалежні, і існує дисперсія кожної величини, то дисперсія суми цих величин дорівнює сумі дисперсій доданків.

Доведення. Нехай маємо дві величини X, Y на ймовірнісному просторі Ω і існують

$$\sigma_1^2 = Var(X) = \sum_{i \geq 1} x_i^2 \Pr(X = x_i) - \left(\sum_{i \geq 1} x_i \Pr(X = x_i) \right)^2,$$

$$\sigma_2^2 = Var(Y) = \sum_{j \geq 1} y_j^2 \Pr(Y = y_j) - \left(\sum_{j \geq 1} y_j \Pr(Y = y_j) \right)^2.$$

Тоді

$$\begin{aligned} Var(X + Y) &= \\ &= \sum_{i,j \geq 1} (x_i + y_j)^2 \Pr(X = x_i, Y = y_j) - \left(\sum_{i \geq 1} x_i \Pr(X = x_i) + \sum_{j \geq 1} y_j \Pr(Y = y_j) \right)^2 \\ &= \sum_{i \geq 1} x_i^2 \Pr(X = x_i) - \left(\sum_{i \geq 1} x_i \Pr(X = x_i) \right)^2 + \sum_{j \geq 1} y_j^2 \Pr(Y = y_j) - \left(\sum_{j \geq 1} y_j \Pr(Y = y_j) \right)^2 \\ &= \sigma_1^2 + \sigma_2^2. \end{aligned}$$

■

Приклад 5.6 Знайдемо гіпергеометричний розподіл ймовірностей $x(n)$ випадкової величини n — того, що в партії із 100 деталей з 10-ма бракованими серед 15 вибраних деталей виявиться рівно n бракованих деталей.

$x(n)$ шукаємо згідно з формулою (15) з потрібними нам параметрами:

$$x(n) = \frac{C_{10}^n \cdot C_{90}^{15-n}}{C_{100}^{15}}.$$

Коли $n < 0$, і коли $n > 10$, тоді $x(n) = 0$. В решті випадків проведемо обчислення:

$$x(0) = 0.1807687, \quad x(1) = 0.3567803, \quad x(2) = 0.2919112, \quad x(3) = 0.129738 \\ x(4) = 0.03448, \quad x(5) = 0.005690, \quad x(6) \sim x(7) \sim x(8) \sim x(9) \sim x(10) \sim 0.$$

Мода випадкової величини, це те значення величини, ймовірність якого найбільша. В нашому випадку найімовірнішим значенням $n \in 1$. Отже модою є 1.

Знайдемо математичне очікування $E(n)$ введеної випадкової величини n , використовуючи уже знайдені значення $x(1), x(2), \dots$:

$$\mu = E(n) = \sum_{i=0}^{10} i \cdot x(i) = 1.5 \dots$$

Дисперсією випадкової величини n

$$\sum_{i=0}^{10} x(i) \cdot (i - \mu)^2 = \sum_{i=0}^{10} x(i) \cdot (i - 1.5)^2 = 1.02 \dots,$$

відхилення — це квадратний корінь із дисперсії, в нашому випадку відхилення з точністю до двох знаків дорівнює 1,01.

6 Загальна алгебра і дискретна ймовірність

6.1 Кільце формальних степеневих рядів

6.1.1 Означення кільця

Кільцем називають непорожню множину M , на якій визначено дві двомісні (бінарні) операції — додавання $a + b$ та множення $a \cdot b$ так, що виконуються наступні аксіоми:

1. відносно додавання кільце є комутативною групою, тобто

- (a) додавання асоціативне — $(a + b) + c = a + (b + c)$ для будь-яких $a, b, c \in M$;
- (b) додавання комутативне — $a + b = b + a$ для будь-яких $a, b \in M$;
- (c) існує 0 — $0 + a = a + 0 = a$ для будь-якого $a \in M$;
- (d) у кожного елемента $a \in M$ існує протилежний $-a$ такий, що $a + (-a) = (-a) + a = 0$.

2. відносно множення кільце є напівгрупою — тобто

- (a) множення асоціативне — $(ab)c = a(bc)$ для будь-яких $a, b, c \in M$;

3. між собою додавання та множення зв'язані двома дистрибутивними законами

- (a) $a(b + c) = ab + ac$;
- (b) $(a + b)c = ac + bc$.

Якщо множення в кільці комутативне, тобто $ab = ba$ для будь-яких $a, b \in M$, то кільце називають комутативним.

Якщо в кільці існує одиниця — $1a = a1 = a$ для будь-якого $a \in M$, то кільце називають кільцем з одиницею.

Приклади кілець: кільце цілих чисел, кільце класів лишок за певним модулем; кільце поліномів; кільце формальних степеневих рядів, кільце квадратних матриць над певним кільцем

У кільці можуть бути оборотні елементи, тобто елементи $a \in M$, для яких існує $b \in M$ такий, що $ab = ba = 1$ — такий елемент $b \in M$ позначають a^{-1} або $\frac{1}{a}$.

Визначення 6.1 *Формальним степеневим рядом від змінної x з коефіцієнтами із кільця R (або, іншими*

словами, над кільцем R , над кільцем коефіцієнтів R) є вираз вигляду

$$a_0 + a_1x + a_2x^2 + \dots = \sum_{i=0}^{\infty} a_i x^i, \quad a_i \in R, \quad i = 0, 1, \dots$$

Числа $a_i \in R$ називають коефіцієнтами ряду. Коефіцієнт a_0 називають вільним членом. За домовленість вважається $x^0 = 1$, $a_0 x^0 = a_0$.

Два ряди збігаються тоді і тільки тоді (за визначенням), коли у них однакові відповідні коефіцієнти.

Доданки ряду з нульовими коефіцієнтами можна не писати. Якщо всі коефіцієнти дорівнюють нулю, то ряд позначають символом 0.

Приклад 6.1 Запис (вираз)

$$2 + 3x + 4x^2 + 5x^3 + \dots = \sum_{n=0}^{\infty} (n+2)x^n$$

є формальним степеневим рядом з вільним членом 2. Коефіцієнтом при x^{24} є 26.

Приклад 6.2 Запис (вираз)

$$2 + 3x + 4x^2 + 0x^3 + 0x^4 + \dots = \sum_{n=0}^{\infty} a_n x^n, \quad a_0 = 2, a_1 = 3, a_2 = 4, \quad a_n = 0 \text{ при } n$$

є формальним степеневим рядом з вільним членом 2 і лише трьома ненульовими доданками.

Визначення 6.2 Поліномами (чи многочленами) називають формальні степеневі ряди, що мають лише скінченну кількість ненульових доданків (коефіцієнтів).

Многочлен $f(x)$

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_n x^n = \sum_{i=0}^n a_i x^i, \quad a_0, a_1, \dots, a_n \in R, \quad a_n \neq 0$$

має вільний член a_0 , старший член $a_n x^n$, старший коефіцієнт $a_n \neq 0$ і степінь $\deg f = n$. Степенем нульового многочлена вважають $-\infty$ — так зручно у багатьох випадках.

Степенями поліномів $x, x^2 + 1, -5 + 2x^{10}$, відповідно, є числа 1, 2, 10.

Доданки $a_0, a_1 x, a_2 x^2, \dots, a_n x^n$ називають одночленами або мономами.

Приклад 6.3 Вираз

$$1 + 5x + 25x^2 + 125x^3 + \dots + 5^n x^n + \dots = \sum_{i=0}^{\infty} 5^i x^i$$

є степеневим рядом над кільцем цілих чисел з вільним членом 1. Цей ряд не є многочленом, оскільки в ньому нескінченно багато ненульових доданків. $125x^3$ є одним із одночленів цього ряду.

Ряд

$$x^2 - 7x^3 - 12x^7$$

є многочленом над кільцем цілих чисел степеня 7 з нульовим вільним членом, старшим членом (мономом, одночленом, доданком) $-12x^7$ і старшим коефіцієнтом -12.

6.1.2 Дії з рядами та многочленами

На множині формальних степеневих рядів над кільцем визначені операції додавання та множення наступним чином:

$$\sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} b_i x^i = \sum_{i=0}^{\infty} (a_i + b_i) x^i, \quad (50)$$

$$\left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{i=0}^{\infty} b_i x^i \right) = \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i b_j \right) x^k. \quad (51)$$

Теорема 6.1 (про властивості додавання та множення рядів)

Нехай ϵ комутативне кільце з одиницею R . Формальні степеневі ряди над R з операціями додавання та множення, що визначені за допомогою (50), (51) утворюють комутативне кільце з одиницею. Многочлени утворюють підкільце кільця рядів.

Доведення. В цьому переконуємося прямою перевіркою.

Перевіряємо аксіому існування ряду, що протилежний даному.

Беремо довільний ряд $f = \sum_{i=0}^{\infty} a_i x^i$ і розглядаємо ряд $g = \sum_{i=0}^{\infty} (-a_i) x^i$. Елементи $-a_i$ існують в R тому, що R — кільце, а в кільці для кожного елемента ϵ протилежний. Далі додаємо

$$f + g = \sum_{i=0}^{\infty} (a_i + (-a_i)) x^i = \sum_{i=0}^{\infty} 0 x^i = 0.$$

Отже $g = -f$ і доведення існування протилежного ряду ми довели (конструктивно, ми точно вказали цей протилежний ряд).

Перевіряємо асоціативність множення рядів. Нехай ϵ три ряди

$$f = \sum_{i=0}^{\infty} a_i x^i, \quad g = \sum_{i=0}^{\infty} b_i x^i, \quad h = \sum_{i=0}^{\infty} c_i x^i,$$

Потрібно довести рівність $(fg)h = f(gh)$. Позначимо $fg = U$, $gh = V$, $S = Uh$, $T = fV$ і

$$U = \sum_{i=0}^{\infty} u_i x^i, \quad V = \sum_{i=0}^{\infty} v_i x^i, \quad S = \sum_{i=0}^{\infty} s_i x^i, \quad T = \sum_{i=0}^{\infty} t_i x^i,$$

Два ряди збігаються тоді і тільки тоді, коли у них однакові коефіцієнти (за визначенням). Отже, нам потрібно

довести рівність $T = S$ і ми це доведемо, коли перевіримо рівність $t_i = s_i$ для будь-якого $i = 0, 1, 2, \dots$

Беремо певне натуральне число чи 0, позначаємо його через i і обчислюємо (знаючи визначення множення рядів) коефіцієнти s_i, t_i .

$$s_i = \sum_{\varepsilon+\gamma=i} u_\varepsilon c_\gamma, \quad t_i = \sum_{\alpha+\delta=i} a_\alpha v_\delta.$$

Подібним чином

$$u_\varepsilon = \sum_{\alpha+\beta=\varepsilon} a_\alpha b_\beta, \quad v_\delta = \sum_{\beta+\gamma=\delta} b_\beta c_\gamma,$$

Тепер підставляємо значення u_ε та v_δ у вирази для s_i, t_i .

$$\begin{aligned} s_i &= \sum_{\varepsilon+\gamma=i} \left(\sum_{\alpha+\beta=\varepsilon} a_\alpha b_\beta \right) c_\gamma = \sum_{\alpha+\beta+\gamma=i} (a_\alpha b_\beta) c_\gamma, \\ t_i &= \sum_{\alpha+\delta=i} a_\alpha \left(\sum_{\beta+\gamma=\delta} b_\beta c_\gamma \right) = \sum_{\alpha+\beta+\gamma=i} a_\alpha (b_\beta c_\gamma) \end{aligned}$$

Далі користуємося тим, що в кільці коефіцієнтів множення асоціативне і при будь-яких α, β, γ виконується рівність

$$(a_\alpha b_\beta) c_\gamma = a_\alpha (b_\beta c_\gamma).$$

Отже, всі доданки в сумі для s_i збігаються із відповідними доданками в сумі для t_i , суми збігаються і доведення рівності $s_i = t_i$ закінчене. Тому закінчене також доведення асоціативності множення рядів.

Подібним чином перевіряють виконання решти аксіом кільця.

■

Приклад 6.4 Візьмемо ряди

$$f = \sum_{i=0}^{\infty} x^i, \quad g = \sum_{i=0}^{\infty} (-1)^i x^i, \quad h = \sum_{i=0}^{\infty} (2i-3)x^i$$

i обчислимо f^2, fh, fg, gh, fgh .

$$f^2 = (1+x+x^2+\dots)(1+x+x^2+\dots) = 1+2x+3x^2+\dots = \sum_{i=0}^{\infty} (i+1)x^i.$$

$$fh = \sum_{i=0}^{\infty} \alpha_i x^i, \quad \alpha_i = \sum_{p+q=i} (2p-3) = \sum_{p=0}^i (2p-3) = i(i+1)-3(i+1) = i^2-2i-3$$

$$fh = \sum_{i=0}^{\infty} (i^2 - 2i - 3)x^i.$$

$$gh = \sum_{i=0}^{\infty} \beta_i x^i, \quad \beta_i = \sum_{p+q=i} (2p-3)(-1)^q = \sum_{p=0}^i (2p-3)(-1)^{i-p} =$$

$$(-1)^i \sum_{p=0}^i (2p-3)(-1)^p = i+1-2(1+(-1)^i),$$

$$gh = \sum_{i=0}^{\infty} (i-1-(-1)^i \cdot 2)x^i.$$

$$fg = \sum_{i=0}^{\infty} \gamma_i x^i, \quad \gamma_i = \sum_{p+q=i} (-1)^q = \frac{1+(-1)^i}{2}, \quad fg = \sum_{i=0}^{\infty} \frac{1+(-1)^i}{2} x^i.$$

$$f \cdot (gh) = \sum_{i=0}^{\infty} \delta_i x^i, \quad \delta_i = \sum_{p+q=i} (p-1-2 \cdot (-1)^p) = \sum_{p=0}^i (p-1-2 \cdot (-1)^p).$$

Оскільки

$$\sum_{p=0}^i p = \frac{1}{2}i(i+1), \quad \sum_{p=0}^i (-1)^p = -1(i+1), \quad \sum_{p=0}^i (-2) \cdot (-1)^p = \begin{cases} 0, & \text{якщо } i \text{ парне} \\ -2, & \text{якщо } i \text{ не парне} \end{cases}$$

то

$$\delta_i = \frac{1}{2}i(i+1) - (i+1) - (1 + (-1)^i),$$

тобто

$$f \cdot (gh) = \sum_{i=0}^{\infty} \left(\frac{1}{2}i(i+1) - (i+1) - (1 + (-1)^i) \right) x^i. \quad (52)$$

Для перевірки проведених обчислень підрахуємо δ_i при малих значеннях i — наприклад, при $i = 0, 1, 2$, двома способами. Один спосіб нам дає формула (52) для обчислення δ_i , а другий спосіб — це пряме перемноження рядів без використання знаків суми.

Отже за першим способом

$$\delta_0 = -3, \quad \delta_1 = -1, \quad \delta_2 = -2.$$

За другим способом

$$f = 1 + x + x^2 + \dots, \quad g = 1 - x + x^2 + \dots \quad h = -3 - x + x^2 + \dots$$

$$gh = -3 + 2x - x^2 + \dots \quad f(gh) = -3 - x - 2x^2 + \dots$$

Визначення 6.3 Два елементи комутативного кільця з одиницею називаються взаємно оберненими (оберненими один до другого), якщо їх добуток дорівнює одиниці.

Деякі ряди в кільці рядів мають обернений ряд. Наприклад,

$$(1 + x + x^2 + x^3 + \dots + x^n + \dots)(1 - x) = 1$$

і ряди $1 - x$ та $\sum_{i=0}^{\infty} x^i$ взаємно обернені.

Теорема 6.2 (про існування оберненого ряду) . Ряд має обернений тоді і тільки тоді, коли його вільний член має обернений в кільці коефіцієнтів.

Доведення. Нехай є ряд $f(x) = \sum_{i=0}^{\infty} a_i x^i$ з коефіцієнтами із кільця R : $a_i \in R$, $i = 0, 1, 2, \dots$

Якщо ряд $g(x) = \sum_{i=0}^{\infty} b_i x^i$ обернений до ряду $f(x)$, тобто $f(x) \cdot g(x) = 1$, і, відповідно,

$$\begin{aligned} a_0 b_0 &= 1, \\ a_0 b_1 + a_1 b_0 &= 0, \\ a_0 b_2 + a_1 b_1 + a_2 b_0 &= 0, \\ &\dots \\ a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \dots + a_n b_0 &= 0, \\ &\dots \end{aligned}$$

то елементи a_0, b_0 взаємно обернені в кільці. Отже a_0 має обернений елемент в кільці коефіцієнтів.

Якщо ж елемент a_0 має обернений елемент в кільці коефіцієнтів, то виписана система рівнянь для знаходження коефіцієнтів ряду $g(x)$ має розв'язок і, відповідно, ряд має обернений.

Теорема доведена повністю. ■

Від рядів можна брати похідну.

Визначення 6.4 *Похідною від ряду $f(x) = \sum_{i=0}^{\infty} a_i x^i$ з коефіцієнтами із кільця R : $a_i \in R$, $i = 0, 1, 2, \dots$ називають ряд $f'(x) = \sum_{i=0}^{\infty} a_{i+1}(i+1)x^i$. Похідні більш високих порядків визначаються індуктивно.*

Потрібно мати на увазі, що коли $a_n \in R$, і k — натуральне число, то під добутком $k \cdot a_n$ розуміють суму k доданків, кожен з яких дорівнює a_n . Тому $k \cdot a_n \in R$.

Приклад 6.5 *Для прикладу візьмемо многочлени $f_1(x) = x + 2x^2 + x^5 + 2x^6 + x^7$, $f_2(x) = x^3$ над полем $\mathbb{Z}_3$. В полі*

$\mathbb{Z}_3$ $1 + 1 + 1 = 0$. Тому

$$f'_1(x) = 1 + 4x + 5x^4 + 12x^5 + 7x^6 = 1 + x + 2x^4 + x^6, \quad f'_2 = 0.$$

Похідна від рядів має ті ж властивості, що і похідна від функції дійсного аргумента:

Теорема 6.3

$$(f + g)' = f' + g', \quad (fg)' = f'g + g'f.$$

Доведення. Доведення здійснюється прямою перевіркою і проводити його не будемо. ■

Кільце многочленів над кільцем R від змінної x будемо позначати $R[x]$. Кільце формальних степеневих рядів над кільцем R від змінної x будемо позначати $R[[x]]$.

6.1.3 Породжуюча функція розподілу — генератриса

В цьому розділі розглядається випадок величини, що приймає лише цілі невід'ємні значення.

Породжуючою функцією випадкової величини (ПФВ) X , що приймає значення $k = 0, 1, 2, \dots$ називається формальний степеневий ряд

$$G(z) = \sum_{k \geq 0} \Pr(X = k) z^k = \sum_{\omega \in \Omega} \Pr(\omega) z^{X(\omega)}. \quad (53)$$

Зрозуміло, що коли величина приймає лише скінченну кількість значень, то ряд (53) є многочленом. Породжуючу функцію ми уже використовували на стор. 29

Для деяких значень z генератрису $G(z)$ можна обчислити — коли виникають нескінченні суми, то використовуються домовленості, що розглянуті на стор. 65. Значення $G(1)$ завжди існує, це сума ймовірностей елементарних подій, тому $G(1) = 1$.

Приклад 6.6 *Породжуючою функцією розподілу ймовірності із прикладу 5.6 буде (нехтуючи знаками після 4-го)*

$$f(z) = 0.1807 + 0.3567 \cdot z + 0.2919 \cdot z^2 + 0.1297 \cdot z^3 + 0.0344 \cdot z^4 + 0.0056 \cdot z^5.$$

Для перевірки того, що точність обчислень прийнятна, можна підрахувати $f(1)$ — це число повинно дорівнювати одиниці:

$$f(1) = 0.1807 + 0.3567 + 0.2919 + 0.1297 + 0.0344 + 0.0056 = 0.999 \neq 1.$$

Те, що сума не дорівнює одиниці, викликано нашим нехтуванням знаків після 4-го в десятковому записі коефіцієнтів.

Приклад 6.7 *Якщо величина має біномний розподіл, тобто ймовірність значення k дорівнює $C_n^k p^k q^{n-k}$, $p + q = 1, p, q \geq 0$, то породжуючою функцією цієї величини буде*

$$(pz + q)^n,$$

оскільки

$$C_0 q^n + C_n^1 q^{n-1} p + \dots + C_n^n p^n = (q + pz)^n.$$

Приклад 6.8 *Якщо величина має геометричний розподіл, тобто ймовірність значення k дорівнює pq^k , $p + q = 1, p, q \geq 0$, то породжуючою функцією цієї величини буде*

$$\frac{p}{1 - qz},$$

оскільки

$$(1 - qz) \cdot \left(\sum_{k \geq 0} pq^k z^k \right) = p.$$

Теорема 6.4 *Нехай $G(z)$ є породжуючою функцією випадкової величини X , яка приймає натуральні значення або 0 і в якій є математичне очікування $E(X)$. Тоді*

$$E(X) = G'(1).$$

Доведення. Для доведення візьмемо породжуючу функцію у вигляді

$$G(z) = \sum_{\omega \in \Omega} \text{Pr}(\omega) z^{X(\omega)},$$

знайдемо її похідну

$$G'(z) = \sum_{\omega \in \Omega} \text{Pr}(\omega) X(\omega) z^{X(\omega)-1},$$

і підставимо значення $z = 1$. Одержимо

$$G'(1) = \sum_{\omega \in \Omega} \text{Pr}(\omega) X(\omega) = E(X).$$

■

Теорема 6.4 дозволяє швидко знаходити математичне очікування величин, для яких відома породжуюча функція.

Приклад 6.9 *Породжуюча функція біномного розподілу $(q+pz)^n$. Отже для знаходження математичного очікування знаходимо похідну*

$$n(q + pz) \cdot p,$$

підставляємо 1 замість z , і одержуємо

$$E(X) = n(p + q)p = np.$$

Приклад 6.10 *Породжуюча функція геометричного розподілу $G(z) = \frac{p}{1-qz}$. Отже для знаходження математичного очікування знаходимо похідну*

$$G'(z) = \frac{pq}{(1 - qz)^2},$$

підставляємо 1 замість z , і одержуємо

$$E(X) = \frac{pq}{1 - q}.$$

Теорема 6.5 Нехай задана величина X , у якої існує $E(X^2)$, і яка має породжуючу функцію $G(z)$. Тоді

$$\text{Var}(X) = G''(1) + G'(1) - (G'(1))^2.$$

Доведення. Доведення полягає в прямому обчисленні:

$$\begin{aligned} E(X^2) &= \sum_{k \geq 0} k^2 \Pr(X = k) = \\ &= \sum_{k \geq 0} (k \cdot (k-1) + k) \Pr(X = k) = \\ &= \sum_{k \geq 0} (k \cdot (k-1) \Pr(X = k) + k \Pr(X = k)) = \\ &= \sum_{k \geq 0} ((k \cdot (k-1) \Pr(X = k) z^{k-2}|_{z=1} + k \Pr(X = k) z^{k-1}|_{z=1})) = \\ &= G''(z)|_{z=1} + G'(z)|_{z=1} = G''(1) + G'(1). \end{aligned}$$

оскільки $E(X) = G'(1)$, то

$$\text{Var}(X) = E(X^2) - (E(X))^2 = G''(1) + G'(1) - (G'(1))^2.$$

■

Приклад 6.11 Породжуюча функція біномного розподілу $(q + pz)^n$. Отже для знаходження математичного очікування знаходимо похідну

$$n(q + pz) \cdot p,$$

підставляємо 1 замість z , і одержуємо

$$E(X) = n(p + q)p = np.$$

Приклад 6.12 Породжуюча функція геометричного розподілу $G(z) = \frac{p}{1-qz}$. Отже для знаходження дисперсії знаходимо першу і другу похідні

$$G'(z) = \frac{pq}{(1-qz)^2}, \quad G''(z) = \frac{pq^2}{(1-q)^3},$$

обчислюємо значення при $z = 1$:

$$G''(1) = \frac{2pq^2}{(1-q)^3} = \frac{2pq}{p^3} = \frac{2q^2}{p^2}, \quad G'(1) = \frac{pq}{(1-q)^2} = \frac{pq}{p^2} = \frac{q}{p},$$

і одержуємо

$$\text{Var}(X) = \frac{2q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q^2 + pq}{p^2} = \frac{q}{p}.$$

6.2 Напівгрупи

6.2.1 Вільна напівгрупа над заданим алфавітом

В дискретній математиці слово “алфавіт” використовується як синонім слова “множина” в тих випадках, коли елементи множини використовуються для побудови скінченних послідовностей. Елементи алфавіту називають символами або буквами. Скінченні послідовності букв називають словами, а дописування одного слова до другого називають множенням слів.

Зазвичай алфавіти є кінченними множинами.

Приклад 6.13 Розглянемо алфавіт $\{U, H\}$. Тут U та H є буквами, послідовність $a = UUNUU$ та послідовність $b = HNUUUU$ є словами. Добутком ab цих слів буде

$$ab = UUNUUNHNUUUU,$$

а добутком ba буде слово

$$ba = HNUUUUUUNUU.$$

Для зручності, вводять також порожнє слово, воно грає роль одиниці при множенні, тобто

$$\emptyset \cdot a = a \cdot \emptyset = a$$

для будь-якого слова a

Зрозуміло, що так введене множення слів є асоціативним.

Множина разом із асоціативною бінарною операцією називається напівгрупою. Якщо операція названа множенням, то напівгрупа називається мультиплікативною

Напівгрупа, елементами якої є слова над певним алфавітом, а множення слів збігається з приписуванням одного слова до другого, називають вільною напівгрупою над алфавітом. Порожнє слово грає роль одиниці в цій напівгрупі і позначається 1.

Для потреб дискретної теорії ймовірності потрібні алфавіти, що складаються із елементарних подій. Вважається, що елементарні події ймовірнісного простору Ω розташовані як звичайні букви алфавіту — одна за одною, в лінійному (алфавітному) порядку:

$$\Omega = \omega_1, \omega_2, \omega_3, \dots$$

Тоді послідовності елементарних подій стають словами, які можна множити. Всі слова також можна розташувати одне за одним, як слова у словниках, в так званому лексикографічному порядку.

6.2.2 Напівгрупове кільце

Нехай є вільна напівгрупа S над заданим алфавітом Ω . В даному мовному оточенні можна уявляти Ω як певний простір елементарних подій, скінченний або злічений. Розглядаємо множину $\mathbb{Z}(S)$ формальних сум

$$u = \sum_{s \in S} u(s) \cdot s, \quad (54)$$

де $u(s) \in \mathbb{Z}$ для будь-якого слова $s \in S$. Доданків в сумі (54) з ненульовими коефіцієнтами може бути нескінченно багато. Такі формальні суми можна додавати:

$$\left(\sum_{s \in S} u(s) \cdot s\right) + \left(\sum_{s \in S} v(s) \cdot s\right) = \sum_{s \in S} (u(s) + v(s)) \cdot s.$$

Оскільки кожне слово довжини $n \geq 0$ можна розкласти у добуток двох слів скінченною множиною способів — $n + 1$ -м способом:

$$a_1 a_2 a_3 \dots a_{n-1} a_n = \emptyset \cdot a_1 \dots a_n = a_1 \cdot a_2 a_3 \dots a_n = a_1 \dots a_{n-1} \cdot a_n = a_1 \dots a_n \cdot \emptyset$$

то формальні суми можна множити множити

$$\left(\sum_{s \in S} u(s) \cdot s\right) \cdot \left(\sum_{s \in S} v(s) \cdot s\right) = \sum_{s \in S} \left(\sum_{s_1, s_2 \in S, s_1 s_2 = s} u(s_1) \cdot v(s_2)\right) \cdot s$$

Множина $\mathbb{Z}(S)$ разом з операціями додавання та множення в ній є кільцем. Вона називається напівгруповим кільцем.

Приклад 6.14 Нехай маємо триелементний простір подій, в якому елементарними подіями є “підкинута монета впала цифрою вверху” (позначаємо цю елементарну подію буквою $\mathcal{C}$); “підкинута монета впала гербом вверху” (позначаємо цю елементарну подію буквою $\mathcal{G}$); “підкинута монета впала на ребро” (позначаємо цю елементарну подію буквою $\mathcal{P}$). Отже $\Omega = \{\mathcal{C}, \mathcal{G}, \mathcal{P}\}$.

В наведеному прикладі послідовності

$$\emptyset, \quad \mathcal{C}, \quad \mathcal{C}\mathcal{C}, \quad \mathcal{P}\mathcal{C}\mathcal{P}\mathcal{R}\mathcal{G}\mathcal{T}\mathcal{T}\mathcal{G}\mathcal{P}\mathcal{C},$$

є елементами напівгрупи $S(\Omega)$.

Суми

$$u_1 = -5 + 3Ц - 15ГРЦЦ, \quad u_2 = 2 - 4ГРЦЦ + Г,$$

є елементами напівгрупового кільця $\mathbb{Z}(S)$ і

$$u_1 + u_2 = -3 + 3Ц - 19ГРЦЦ + Г$$

$$u_1 u_2 = -10 + 6Ц - 10ГРЦЦ - 12ЦГРЦЦ + 60ГРЦЦГРЦЦ - 5Г + 3ЦГ - 15ГРЦЦ$$

$$u_2 u_1 = -10 + 6Ц - 10ГРЦЦ - 12ГРЦЦЦ + 60ГРЦЦГРЦЦ - 5Г + 3ГЦ - 15ГГРЦЦ$$

Слова a, b, c в добутку abc називаються підсловами.

6.2.3 Приклади знаходження генератриси

Приклад 6.15 Підкидаємо монету з двома можливими результатами $У$ та $Н$ до тих пір, поки підряд не відбудуться (вперше з'явиться підслово) $УУН$. Позначимо через α_n ймовірність того, що $УУН$ вперше з'явиться після n -го підкидання. Знайти генератрису

$$G(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \dots = \sum_{k \geq 0} \alpha_k z^k.$$

В прикладі явно мається на увазі, що

$$\Pr(Y) = \Pr(H) = \frac{1}{2},$$

і підкидання незалежні, так що ймовірність появи заданої послідовності $s = \omega_1 \omega_2 \dots \omega_k$ — слова s в алфавіті $\Omega = \{У, Н\}$ дорівнює 2^{-k} .

Ні за порожню множину підкидань, ні за одне підкидання, ні за два потрібне триелементне слово не одержиш. Тому

$$\alpha_0 = \alpha_1 = \alpha_2 = 0.$$

Оскільки порібне триелементне слово єдине, то

$$\alpha_3 = \frac{1}{8}.$$

Чотириелементних послідовностей, які закінчуються на УУН, дві — УУУН та НУУН. Тому $\alpha_4 = \frac{2}{16} = \frac{1}{8}$.

Розглянемо вільну напівгрупу $S(\Omega)$ слів в алфавіті $\Omega = \{У, Н\}$, напівгрупове кільце $\mathbb{Z}(S)$, і два елемента $a, b \in \mathbb{Z}(S)$ в цьому кільці. Нехай a є сума (з коефіцієнтами 1) потрібних нам слів — тих, що закінчуються підсловом УУН, але раніше такого підслова не мають. Через b позначаємо суму тих слів, які взагалі підслово УУН не містять. Отже,

$$a = \text{УУН} + \text{УУУН} + \text{НУУН} + \text{УУУУН} + \text{УНУУН} + \text{НУУУН} + \text{ННУУН} + \dots$$

$$b = 1 + \text{У} + \text{Н} + \text{УУ} + \text{УН} + \text{НУ} + \text{УУУ} + \text{УНУ} + \text{УНН} + \text{ННУ} + \text{НУУ} + \dots$$

Якщо домножити доданок в b на У, то ми знову одержимо доданок в b , що закінчується символом У. І всі доданки в b , які закінчуються У, можна одержати множенням іншого доданка на У. Якщо домножити доданок в b на Н, то ми матимемо два випадки. Перший — це доданок закінчувався на УУ, і тоді після множення на Н ми одержимо доданок із a . Другий випадок — доданок не закінчується на УУ, тоді ми одержимо доданок із b , який закінчується символом Н. Такими міркуваннями одержимо систему рівнянь

$$\begin{cases} b(У + Н) = b + a - 1, \\ b \cdot \text{УУН} = a. \end{cases} \quad (55)$$

Якщо в рівняння (56) підставити $\frac{1}{2}z$ замість У і замість Н, то сума a перетвориться в потрібну нам породжуючу функцію $G(z)$, b перетвориться в певний степневий ряд $H(z)$, і ми одержимо систему рівнянь

$$\begin{cases} H(z)(\frac{1}{2}z + \frac{1}{2}z) = H(z) + G(z) - 1, \\ H(z) \cdot \frac{1}{8}z^3 = G(z). \end{cases} \quad (56)$$

Розв'язуємо цю систему рівнянь і знаходимо

$$G(z) = \frac{z^3}{z^3 - 8(z - 1)}.$$

Оскільки $G(1) = 1$, то ймовірність того, що послідовність підкидань монети не закінчиться ніколи, дорівнює 0.

Приклад 6.16 Підкидаємо монету з двома можливими результатами $У$ та $Н$ до тих пір, поки підряд не відбудуться $УНН$. Позначимо через α_n ймовірність того, що $УНН$ вперше з'явиться на n -му підкиданні. Знайти генератрису

$$G(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \dots = \sum_{k \geq 0} \alpha_k z^k.$$

Цей приклад відрізняється від попереднього лише фінішуючою послідовністю — там була $УУН$ а тут $УНН$. Ті ж міркування, що і в попередньому прикладі приводять до тієї ж породуючої функції.

Приклад 6.17 Підкидаємо монету з двома можливими результатами $У$ та $Н$ до тих пір, поки підряд не з'явиться підслово $УУН$ або $УНН$. Позначимо через α ймовірність того, що $УУН$ з'явиться раніше від $УНН$, а через β ймовірність того, що раніше з'явиться $УНН$. Знайдемо ці дві ймовірності.

Подібно до попередніх двох прикладів, створюємо три суми $a, b, c \in \mathbb{Z}(S(\Omega))$ слів:

a — сума слів, які закінчуються першою появою $УУН$ ($УНН$ — не з'являється);

b — сума слів, які закінчуються першою появою $УНН$ ($УУН$ — не з'являється);

c — сума слів, що не містять ні $УУН$ ні $УНН$.

і виписуємо для цих сум рівняння

$$1 + c(Y + H) = c + a + b;$$

$$cYUH = a;$$

$$cYHH = aH + b.$$

В одержані три рівняння підставляємо $\frac{1}{2}$ замість Y та H . При цьому a перейде в α , b перейде в β і c перейде в деяке число γ . Для цих чисел маємо рівняння

$$1 + \gamma = \gamma + \alpha + \beta;$$

$$\frac{1}{8}\gamma = \alpha,$$

$$\frac{1}{8}\gamma = \frac{1}{2}\alpha + \beta.$$

Розв'язавши ці рівняння одержуємо відповіді

$$\alpha = \frac{2}{3}, \quad \beta = \frac{1}{3}.$$

На перший погляд одержаний результат, а саме $\alpha \neq \beta$, є дивним, тому що і ймовірність послідовності УУН і ймовірність послідовності УНН одна і та ж, вона дорівнює $\frac{1}{8}$.

6.2.4 Напівгрупа стохастичних матриць і ланцюги Маркова

Перед тим, як вести нові поняття, розглянути нові підходи до знаходження ймовірностей, розберемо конкретний приклад.

Приклад 6.18 Припустимо, що Лариса обслуговує об'їздування в Ковелі, Миргороді і в Бахмачі, і, відповідно, кожен місяць має відрядження або в Ковель, або в Миргород, або в Бахмач. Ймовірність знаходження Лариси у відрядженні в певному місяці залежить від

того, де вона була в попередній місяць. Якщо вона була в Ковелі, то ймовірності бути в наступному місяці в Ковелі, Миргороді чи в Бахмачі дорівнюють, відповідно, $0, \frac{1}{2}, \frac{1}{2}$; якщо була в Миргороді, то ймовірності бути в наступному місяці в Ковелі, Миргороді чи в Бахмачі дорівнюють, відповідно, $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$; і якщо була в Бахмачі, то відповідні ймовірності дорівнюють $\frac{3}{7}, \frac{2}{7}, \frac{2}{7}$. Якщо на даний момент часу Лариса знаходиться в Ковелі, яка ймовірність того, що вона буде через 12 місяців у відрядженні в Ковелі, Миргороді чи в Бахмачі?

Спочатку введемо ймовірнісний простір — простір елементарних подій. Їх три

K — “Лариса знаходиться в Ковелі”;

M — “Лариса знаходиться в Миргороді”;

B — “Лариса знаходиться в Бахмачі”.

Зараз маємо елементарну подію a_0 , $a_0 = K$. Далі йде послідовність із 12 елементарних подій

$$a = a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} a_{11} a_{12}.$$

Потрібно знайти ймовірності

$$p_{12} = \Pr(a_{12} = K), \quad q_{12} = \Pr(a_{12} = M), \quad r_{12} = \Pr(a_{12} = B).$$

Елементарні події K, M, B не мають ймовірностей, немає функції розподілу ймовірності. Відомо лише, що

$$\Pr(a_0 = K) = 1, \quad \Pr(a_0 = M) = 0, \quad \Pr(a_0 = B) = 0,$$

і відомі умовні ймовірності — ймовірності того, що Лариса знаходиться в певному місті в певному місяці, за умови, що відомо, де вона була в попередній місяць:

$$\Pr(K|K) = 0, \quad \Pr(K|M) = \frac{1}{3}, \quad \Pr(K|B) = \frac{3}{7},$$

$$\begin{aligned}\Pr(M|K) &= \frac{1}{2}, & \Pr(M|M) &= \frac{1}{3}, & \Pr(M|B) &= \frac{2}{7}, \\ \Pr(B|K) &= \frac{1}{2}, & \Pr(B|M) &= \frac{1}{3}, & \Pr(B|B) &= \frac{2}{7}.\end{aligned}$$

Запишемо множину відомих ймовірностей у вигляді матриці

$$A = \begin{pmatrix} \Pr(K|K) & \Pr(M|K) & \Pr(B|K) \\ \Pr(K|M) & \Pr(M|M) & \Pr(B|M) \\ \Pr(K|B) & \Pr(M|B) & \Pr(B|B) \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{3}{7} & \frac{2}{7} & \frac{2}{7} \end{pmatrix}.$$

Припустимо, що Лариса в певний місяць знаходиться в містах Ковель, Миргород чи Бахмач з імовірностями p, q, r . Тоді за формулою повної ймовірності p', q', r' ймовірності того, що Лариса в наступний місяць знаходиться в цих містах, буде дорівнювати

$$\begin{aligned}p' &= p \cdot \Pr(K|K) + q \cdot \Pr(K|M) + r \cdot \Pr(K|B), \\ q' &= p \cdot \Pr(M|K) + q \cdot \Pr(M|M) + r \cdot \Pr(M|B), \\ r' &= p \cdot \Pr(B|K) + q \cdot \Pr(B|M) + r \cdot \Pr(B|B).\end{aligned}$$

Запишемо останні три рівності в матричному вигляді:

$$(p', q', r') = (p, q, r) \cdot A.$$

Остання рівність дозволяє записати

$$(p_{12}, q_{12}, r_{12}) = (1, 0, 0) \cdot A^{12}$$

За допомогою пакета символьних обчислень знайдемо A^{12} :

$$A^{12} = \begin{pmatrix} \frac{4601050272104}{16679880978201} & \frac{12078830706097}{33359761956402} & \frac{12078830706097}{33359761956402} \\ \frac{13804224558136}{50039642934603} & \frac{36235418376467}{100079285869206} & \frac{36235418376467}{100079285869206} \\ \frac{32210573130200}{116759166847407} & \frac{84548593717207}{233518333694814} & \frac{84548593717207}{233518333694814} \end{pmatrix},$$

або з точністю до трьох знаків після коми

$$A^{12} \approx \begin{pmatrix} 0,276 & 0,362 & 0,362 \\ 0,276 & 0,362 & 0,362 \\ 0,276 & 0,362 & 0,362 \end{pmatrix}$$

Отже

$$p_{12} = 0,276, \quad q_{12} = 0,362, \quad r_{12} = 0,362.$$

Виділимо кілька особливостей прикладу, що розглядався.

Перша особливість полягає в тому, що ймовірність елементарної події (Лариса знаходиться в певному місці) в ланцюгу подій залежала від того, якою була ймовірність попередньої елементарної події (де була Лариса в попередній час). Такі ланцюги подій називають ланцюгами Маркова.

Друга особливість полягає в тому, що ми використали матрицю, елементами якої є певні ймовірності, — рядками матриці були розподіли ймовірності елементарних подій за умови, що відбулася певна попередня подія. Такі матриці називають ймовірнісними або стохастичними. Отже

Квадратна матриця називається стохастичною, якщо її елементи — невід’ємні дійсні числа і сума всіх елементів кожного рядка дорівнює 1.

Теорема 6.6 *Стохастичні матриці даного розміру разом з операцією множення матриць утворюють напівгрупу.*

Доведення. Нагадаємо, що напівгрупою називають множину з асоціативною бінарною операцією. Операція множення матриць асоціативна, це відомо з курсу алгебри. Тому для доведення теореми потрібно лише пе-

ресвідчитися, що добуток двох стохастичних матриць знову є стохастичною матрицею.

Нехай

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & & & \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

— дві стохастичні матриці і $C = A \cdot B$ є добутком цих матриць. Потрібно довести, що матриця C є стохастичною, тобто для будь-якого $i = 1, 2, \dots, n$ буде

$$u = \sum_{j=1}^n c_{ij} = 1.$$

За означенням добутку двох матриць

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

Тому

$$u = \sum_{j=1}^n \left(\sum_{k=1}^n a_{ik} b_{kj} \right).$$

змінивши порядок підсумовування одержимо

$$u = \sum_{k=1}^n \left(\sum_{j=1}^n a_{ik} b_{kj} \right) = \sum_{k=1}^n a_{ik} \left(\sum_{j=1}^n b_{kj} \right).$$

За умовою, матриця B стохастична і при будь-якому $k = 1, 2, \dots, n$ $\sum_{j=1}^n b_{kj} = 1$. Отже

$$u = \sum_{k=1}^n a_{ik},$$

а права частина цієї рівності дорівнює 1 тому, що матриця A стохастична.

Теорема доведена. ■

Стохастичні матриці в мовному оточенні ланцюгів Маркова називають також “матрицями переходів”, а самі елементарні події називають станами — станами марківського процесу.

Приклад 6.19 *Припустимо, що система може бути в станах 1,2,3,4 і кожен день випадково змінює свій стан — всі переходи рівноймовірні і не залежать від того, в якому стані система була перед переходом, тобто матриця переходів має вигляд*

$$A = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}.$$

На сьогодні система знаходиться в стані 1. Завтра вона буде в стані 4 з ймовірністю 0,25. Яка ймовірність того, що за наступні 5 днів система хоч один раз попаде в стан 4?

Оскільки нас цікавить попадання в 4 стан, то змінимо ймовірність переходів із 4-го стану — нехай система залишається в 4-му стані, якщо вже туди попала. Так що матриця переходів прийме вигляд

$$A = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Підраховуємо, знову за допомогою пакета символьних обчислень, A^5 :

$$A^5 = \begin{pmatrix} 0.0791015625 & 0.0791015625 & 0.0791015625 & 0.7626953125 \\ 0.0791015625 & 0.0791015625 & 0.0791015625 & 0.7626953125 \\ 0.0791015625 & 0.0791015625 & 0.0791015625 & 0.7626953125 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{pmatrix}$$

З точністю до двох знаків

$$(1, 0, 0, 0) \cdot A^5 = (0.08, 0.08, 0.08, 0.76).$$

Тому за 5 днів система бодай один раз попаде в стан 4 з імовірністю 0,76.

6.3 Рекурентні послідовності.

6.3.1 Процеси і послідовності

Латинське слово *processus* означає послідовну зміну, розвиток в часі, предметів, явищ, властивостей предметів та подібного. Залежність процесу (точніше стану процесу) від часу є суттєвою властивістю процесу. Оскільки ми знайомимося з дискретною теорією ймовірностей, то ми розглядаємо дискретні процеси, тобто такі, в яких

- значень часу або скінченна кількість (процес відбувається за скінченну кількість кроків), або зліченна, тобто значення часу, в які визначається стан процесу, можна розташувати у послідовність — час дискретний;
- стан процесу у відмічений момент часу є дискретною випадковою подією.

Отже, процес a характеризується моментами часу $t_0, t_1, t_2, t_3, \dots$ і станами $a(t_1), a(t_2), a(t_3), \dots$. Щоб підкреслити, що стан є випадковою подією, який має певну

йомвірність, процес називають стохастичним (від грецького слова *στοχασις*, яке означає здогадку), або випадковим.

В деяких стохастичних процесах звертають увагу не на всі елементарні події, які можуть відбутися, а тільки на певні. Так на телефонну станцію можуть поступати виклики, а можуть не поступати — елементарних подій дві. Але цікавить лише одна — поступання виклику. Страхову компанію цікавить настання страхової події. В таких мовних оточеннях стохастичні процеси називають також потоками випадкових подій.

Сказане вище є мотивацією наступного означення:

Нехай Ω є непорожня генеральна сукупність елементарних подій Ω і $T \subseteq \mathbb{R}$. Відображення

$$f : T \rightarrow \Omega$$

називається процесом.

Одним із способів задання відображення f із скінченної чи зліченної множини є табличний. Таблиця, що задає відображення f , має два рядки — у верхньому рядку стоять можливі значення аргумента x , а в нижньому — відповідні значення функції $f(x)$. Наприклад, функцію x^2 від натурального аргумента можна задати так:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \dots \\ 1 & 4 & 9 & 16 & 25 & \dots \end{pmatrix}$$

Якщо множина аргументів природно впорядкована, то верхній рядок в табличному заданні функції можна не писати, — функція задається лише одним нижнім рядком — послідовністю. Таким чином, кожну послідовність a елементів із Ω

$$a = a_1, a_2, a_3 \dots$$

можна вважати функцією із множини $1, 2, 3, \dots$ у множину Ω , і навпаки, кожен функцію із множини $1, 2, 3, \dots$ у множину Ω можна так задати. Сказане дозволяє в дискретному випадку не рорізняти потоки, процеси і послідовності подій із генеральної сукупності.

6.3.2 Рекурентні співвідношення та послідовності¹².

В даному розділі розглядаємо послідовності

$$u_0, u_1, \dots, u_n, \dots; \quad u_i \in \mathbb{R}. \quad (57)$$

у яких n -й член можна обчислити, коли відомі попередні члени, за допомогою формули

$$u_n = \alpha_1 u_{n-1} + \alpha_2 u_{n-2} + \dots + \alpha_k u_{n-k}, \quad \alpha_k \neq 0; u_i, \alpha_i \in \mathbb{R}, i = 0, 1, 2, \dots \quad (58)$$

Взагалі, коли в послідовності кожен елемент може бути обчислений за певною формулою, при відомих попередніх (звичайно, починаючи з певного місця), тоді таку послідовність називають рекурентною, а формулу, яка дозволяє провести це обчислення — рекурентним співвідношенням. Співвідношення (58) називають лінійним однорідним рекурентним співвідношенням — воно часто зустрічається як в теорії мовірностей, так і в комбінаториці, в інших розділах математики (алгебра, математичний аналіз, диференціальні рівняння, тощо).

Для довільного члена послідовності (57), яка задовольняє лінійному рекурентному співвідношенню (58), можна знайти формулу, яка дозволяє обчислити цей член.

¹²Лінійні рекурентні співвідношення ретельно розглядаються при вивченні визначників матриць і в комбінаториці. Тому наразі текст має характер нагадування

Послідовність $0, 2, 4, 6, \dots$ задовольняє лінійному однорідному рекурентному співвідношенню $u_n = u_{n-1} + 2, n = 1, 2, \dots$. Це арифметрична прогресія. Ми можемо вказати формулу для загального члена цієї послідовності: $u_n = 2n, n = 0, 1, 2, \dots$. Прикладом послідовності, яка задовольняє лінійному однорідному рекурентному співвідношенню, є геометрична прогресія.

Щоб знайти формулу для рекурентної послідовності (57), яка задовольняє лінійному однорідному співвідношенню (58), спочатку шукають відповідь у вигляді геометричної прогресії, потім користуються тим, що сума послідовностей, що задовольняють (58), знову задовольняє (58). І якщо послідовність задовольняє (58), то помножена на будь-яке число вона знову задовольняє цьому лінійному однорідному співвідношенню.

Геометрична послідовність

$$a, aq, aq^2, aq^3, \dots$$

задовольняє лінійному однорідному рекурентному співвідношенню (58) тоді і тільки тоді, коли або $a = 0$, або q є коренем рівняння

$$\chi(x) = x^k - \alpha_1 x^{k-1} - \alpha_2 x^{k-2} - \dots - \alpha_k.$$

Якщо $k = 2, \alpha_2 \neq 0$, то кажуть що співвідношення має глибину 2. В такому випадку послідовність (57) має рекурентне співвідношення

$$u_n = \alpha_1 u_{n-1} + \alpha_2 u_{n-2}, \quad \alpha_2 \neq 0, \quad (59)$$

і загальний член u_n цієї послідовності має вигляд

$$u_n = c_1 x_1^n + c_2 x_2^n, \quad (60)$$

якщо x_1, x_2 є різними коренями рівняння $x^2 - \alpha_1 x - \alpha_2 = 0$, (c_1, c_2 — деякі сталі. Якщо ж останє рівняння має два

однакові корені $x_1 = x_2$, то

$$u_n = (c_1 n + c_2)x_1^n \quad (61)$$

для деяких сталих c_1, c_2 .

6.3.3 Задача про розорення гравця

Можливості методу рекурентних співвідношень виявимо на наступному прикладі.

Приклад 6.20 Два гравці мають сумарний капітал a гривень і грають “в орлянку” — якщо випадає герб (Γ), то перший гравець віддає гривню другому гравцю, а якщо випадає цифра (Π), то він одержує гривню від другого гравця. Припустимо, що гривня падає гербом наверх із ймовірністю p , а падає цифрою вгору із ймовірністю q — $p + q = 1$. Гра припиняється, коли один із гравців збанкрутує — тобто в нього не залишиться грошей для продовження гри.

Яка ймовірність того, що

- перший гравець збанкрутує;
- другий гравець збанкрутує;
- гра не матиме кінця — продовжуватиметься нескінченно довго?

Позначимо через p_n ймовірність того, що перший гравець збанкрутує до $n + 1$ -го підкидання монети. Тоді $p_n \geq p_m$, якщо $n \geq m$, послідовність

$$p_0, p_1, p_2, \dots$$

монотонно не спадає, обмежена 1 і тому має границю

$$\lim_{n \rightarrow \infty} p_n.$$

От цю границю і називаємо ймовірністю того, що перший гравець збанкрутує.

Позначимо через $u(n)$ ймовірність того, що перший гравець збанкрутує маючи початковий капітал n гривень. Оскільки капітал першого гравця після підкидання монети зменшується на 1 з ймовірністю p і збільшується на 1 з ймовірністю q , то для будь-якого $n \geq 1$

$$u(n) = p \cdot u(n-1) + q \cdot u(n+1),$$

або

$$q \cdot u(n+1) - u(n) + p \cdot u(n-1) = 0. \quad (62)$$

А це означає, що послідовність

$$u(0), u(1), u(2), \dots$$

задовольняє лінійному однорідному рекурентному співвідношенню (62) і $u(n)$ можна знайти за формулою

$$u(n) = c_1 x_1^n + c_2 x_2^n,$$

якщо корені x_1, x_2 рівняння

$$q \cdot x^2 - x + p = 0 \quad (63)$$

різні, і за формулою

$$u(n) = (c_1 n + c_2) x_1^n,$$

якщо ці корені однакові.

Оскільки $p + q = 1$, і

$$p - 1 + q = 0,$$

$$q \frac{p^2}{q^2} - \frac{p}{q} + p = \frac{p^2}{q} = \frac{p}{q} + q = \frac{p(1-p)}{q} + p = 0,$$

то коренями рівняння (63) є

$$x_1 = 1, \quad x_2 = \frac{p}{q}.$$

Тому для деяких дійсних чисел c_1, c_2

$$u(n) = c_1 \cdot 1^n + c_2 \left(\frac{p}{q}\right)^n,$$

якщо $p \neq q$, і

$$u(n) = c_1 \cdot n + c_2,$$

якщо $p = q = \frac{1}{2}$.

Сталі c_1, c_2 знаходимо із умов $u(0) = 1$ — якщо до початку гри перший гравець не має грошей, то він уже гарантовано банкрут, і $u(a) = 0$ — якщо перший гравець має весь сумісний капітал, то другий гравець збанкрутував і гра припиняється, перший гравець гарантовано не збанкрутує. Отже, коли $p \neq q$,

$$c_1 + c_2 = 1,$$

$$c_1 + c_2 \left(\frac{p}{q}\right)^a = 0,$$

Розв'язками цієї системи є

$$c_2 = - \left(-1 + \left(\frac{p}{q}\right)^a\right)^{-1}, \quad c_1 = \left(\frac{p}{q}\right)^a \left(-1 + \left(\frac{p}{q}\right)^a\right)^{-1}$$

і можна записати відповідь

$$u(n) = \left(\frac{p}{q}\right)^a \left(-1 + \left(\frac{p}{q}\right)^a\right)^{-1} - \left(-1 + \left(\frac{p}{q}\right)^a\right)^{-1} \left(\frac{p}{q}\right)^n = \frac{\left(\frac{p}{q}\right)^a - \left(\frac{p}{q}\right)^n}{-1 + \left(\frac{p}{q}\right)^a}.$$

Якщо $p = q$, то для знаходження c_1, c_2 маємо рівняння $c_2 = 1$, $c_1 a + c_2 = 0$. В такому випадку $c_1 = -\frac{1}{a}$

$$u(n) = -\frac{1}{a} \cdot n + 1 = \frac{a - n}{a}.$$

Відповідь на перше питання отримана. Точно так же шукаємо відповідь на друге питання.

Позначимо через $v(n)$ ймовірність того, що другий гравець збанкрутує маючи стартовий капітал n . Для послідовності $v(0), v(1), v(2), \dots$ маємо рекурентне співвідношення

$$v(n) = pv(n+1) + qv(n-1), \quad pv(n+1) - v(n) + qv(n-1) = 0.$$

Загальний член цієї послідовності шукаємо у вигляді

$$v(n) = c_1 + c_2 \left(\frac{q}{p}\right)^n$$

якщо $p \neq q$, і

$$v(n) = c_1 n + c_2,$$

якщо $p = q$. Сталі c_1, c_2 знаходимо із умов $v(0) = 1, v(a) = 0$. Одержуємо

$$c_2 = -\left(-1 + \left(\frac{q}{p}\right)^a\right)^{-1}, \quad c_1 = \left(\frac{q}{p}\right)^a \left(-1 + \left(\frac{q}{p}\right)^a\right)^{-1}$$

і можна записати відповідь

$$v(n) = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^n}{-1 + \left(\frac{q}{p}\right)^a},$$

Якщо $p = q$, то для знаходження c_1, c_2 маємо рівняння $c_2 = 1, \quad c_1 a + c_2 = 0$. В такому випадку $c_1 = -\frac{1}{a}$

$$v(n) = -\frac{1}{a} \cdot n + 1 = \frac{a - n}{a}.$$

Відповідь на друге питання отримана. Перейдемо до питання про ймовірність w того, що гра ніколи не припиниться.

Для будь-якого $0 \leq n \leq a$ $u(n) + v(a - n) + w = 1$ і $w = 1 - (u(n) + v(a - n))$.

Оскільки при $p \neq q$

$$u(n) + v(a - n) = \frac{\left(\frac{p}{q}\right)^a - \left(\frac{p}{q}\right)^n}{-1 + \left(\frac{p}{q}\right)^{n-a}} + \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^n}{-1 + \left(\frac{q}{p}\right)^{n-a}} = 1,$$

і при $p = q$ $u(n) + v(a - n) = \frac{a-n}{a} + \frac{n}{a} = 1$, то

$$w = 0,$$

тобто без кінця продовжуватися гра не може.

Література

- [1] Р.Грэхем, Д.Кнут, О.Паташник “Конкретная математика. Основания информатики” — Москва:Мир, 1998.
- [2] В.Феллер “Введение в теорию вероятностей и её приложения”, в 2-х томах. — Москва:Мир. 1967
- [3] В.Е.Гмурман “Теория вероятностей и математическая статистика”. — Москва:Высшая школа, 1972

Показчик

алфавіт, 86	інтегральна, 61
буква, 86	кумулятивна, 61
час	функція
дискретний, 99	породжуюча
частота, 12	розподілу, 83
дисперсія, 70, 73	генератриса, 82
добуток	індукція
подій, 8	математична, 22
порожньої множини мно-	ймовірність, 6
жників, 31	елементарної події, 6
рядів, 77	гіпотези
величин, 63	апостеріорна, 53
додавання	апріорна, 52
формальних сум, 88	події, 6
рядів, 77	величини, 57, 72
доведення	кільце
конструктивне, 77	многочленів, 82
формула	напівгрупове, 88
Байєса, 50	коефіцієнт
Бернуллі, 35	при степені змінної, 75
Пуассона, 24, 35	ряду, 75
повної ймовірності, 46	ланцюг
формули	Маркова, 92, 95
Байєса, 50	математичне очікування
функція	суми величин, 69
породжуюча, 82	математичне очікування, 67
розподілу	добутку
кумулятивна, 63	незалежних величин, 69
накопичувальна, 63	матриця
розподілу ймовірності, 7	переходів, 97
функція розподілу	межа

точна верхня, 6	рівноймовірні, 7
многочлен, 76	несумісні, 8
ступінь, 76	незалежні, 41
множення	повна група, 8
формальних сум, 88	рівноймовірні, 12
рядів, 80	поліном, 76
слів, 86	старший член, 76
множина, 3	ступінь, 76
моном, 76	вільний член, 76
напівгрупа, 87	порядок
мультиплікативна, 87	лексикографічний, 87
об'єднання	лінійний, 87
подій, 8	послідовність
оцінка гіпотези	рекурентна, 100
апостеріорна, 53	потік
апріорна, 52	випадкових подій, 99
очікування	процес, 99
математичне, 73	дискретний, 99
одиниця	марківський, 97
напівгрупи, 87	стохастичний, 99
одночлен, 76	випадковий, 99
параметр	прогресія
розподілу Пуассона, 23	арифметична, 101
перетин	геометрична, 101
подій, 8	простір
підслово, 89	подій, 3
подія, 3	розміщення, 9
елементарна, 3	з повтореннями, 9
неможлива, 7	розподіл
порожня, 7	Бернуллі, 14
сприятлива, 8	Пуассона, 24, 68
події	біномний, 14, 68
елементарні	геометричний, 27

ймовірності, 7	рядів, 77
рівномірний, 7	величин, 63
ймовірності величини, 63	супремум, 6
біномний, 63	теорія
поліноміальний, 32	ймовірностей, 3
рівномірний, 12, 14	величина, 72
від'ємний біномний, 28	випадкова, 56
розподіл ймовірності	величини
величини	незалежні, 63, 69
рівномірний, 68	вибірка, 9
ряд	впорядкована
формальний степеневий,	з поверненнями, 9
74	з поверненнями, 9
обернений до ряду, 80	випробування
протилежний, 77	Бернуллі, 14
слово, 86	відхилення
співвідношення	середнє квадратичне, 70
лінійне	стандартне, 70, 73
глибини 2, 100	
лінійне неоднорідне, 100	
рекурентне, 21, 100	
лінійне однорідне, 100	
сполука, 9	
з повтореннями, 9	
стан	
марківського процесу, 97	
процесу, 99	
сукупність	
генеральна, 3	
сума	
формальна, 87	
нескінченна, 7	
подій, 8	