

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені Ігоря СІКОРСЬКОГО»  
Навчально-науковий фізико-технічний інститут  
Кафедра математичних методів захисту інформації

«На правах рукопису»

УДК (519.2)

«До захисту допущено»

Завідувач кафедри

\_\_\_\_\_ Сергій ЯКОВЛЄВ

«\_\_» \_\_\_\_\_ 2026 р.

**Дипломна робота**  
**на здобуття ступеня бакалавра**  
**за освітньо-професійною програмою**  
**«Математичні методи криптографічного захисту інформації»**

зі спеціальності: 113 Прикладна математика  
на тему: **«Асимптотичний аналіз проріджених випадкових сум»**

Виконала:

студентка IV курсу, групи ФІ-23

Дегтярьова Марія Іванівна \_\_\_\_\_

Керівник:

доц. каф. математичних методів захисту інформації,

кандидат фіз.-мат. наук, доц.

Ніщенко Ірина Іванівна \_\_\_\_\_

Рецензент:

ст. наук. співроб. відділу теорії випадкових процесів

Інституту математики НАН України, кандидат

фіз.-мат. наук

Рябов Георгій Валентинович \_\_\_\_\_

Засвідчую, що у цій дипломній  
роботі немає запозичень з праць  
інших авторів без відповідних  
посилань.

Студентка \_\_\_\_\_

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені Ігоря СІКОРСЬКОГО»

Навчально-науковий фізико-технічний інститут  
Кафедра математичних методів захисту інформації

Рівень вищої освіти — перший (бакалаврський)  
Спеціальність — 113 Прикладна математика,  
ОПП «Математичні методи криптографічного захисту інформації»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Сергій ЯКОВЛЄВ

«\_\_» \_\_\_\_\_ 2026 р.

**ЗАВДАННЯ**  
**на дипломну роботу**

Студентка: Дегтярьова Марія Іванівна

1. Тема роботи: *«Асимптотичний аналіз проріджених випадкових сум»*, науковий керівник дипломної роботи: доц. каф. математичних методів захисту інформації, кандидат фіз.-мат. наук, доц. Ніщенко Ірина Іванівна,

затверджені наказом по університету №\_\_ від «\_\_» \_\_\_\_\_ 2026 р.

2. Термін подання студентом роботи: «\_\_» \_\_\_\_\_ 2026 р.

3. Об'єкт дослідження: випадкові проріджені суми незалежних випадкових величин.

4. Предмет дослідження: асимптотичні властивості проріджених випадкових сум залежно від структури множини індексів.

5. Перелік завдань:

1) розглянути приклади обчислення натуральної щільності множин;  
2) дослідити властивості повних та проріджених випадкових сум;  
3) отримати асимптотичні оцінки різниці між повною та прорідженою сумами;

4) встановити умови збіжності проріджених сум до нормального розподілу;

5) дослідити швидкість збіжності до нормального розподілу та застосувати оцінки типу нерівності Беррі–Ессеєна;

6) проаналізувати вплив параметра, що визначає асимптотичну поведінку множини індексів, на поведінку сум;

7) провести чисельні експерименти та візуалізацію отриманих результатів.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: презентація доповіді

7. Орієнтовний перелік публікацій: основні результати роботи доповідалися на семінарі для студентів з додаткових глав теорії ймовірностей та функціонального аналізу при відділі Інституту математики НАН України та XXIV Всеукраїнській науково-практичній конференції студентів, аспірантів та молодих вчених “Теоретичні і прикладні проблеми фізики, математики та інформатики”.

8. Дата видачі завдання: 10 вересня 2025 р.

## Календарний план

| № з/п | Назва етапів виконання дипломної роботи                                                          | Термін виконання         | Примітка |
|-------|--------------------------------------------------------------------------------------------------|--------------------------|----------|
| 1     | Узгодження теми роботи із науковим керівником                                                    | 01-15 вересня 2025 р.    | Виконано |
| 2     | Огляд літературних джерел та вивчення основних понять натуральної щільності та випадкових сум    | Вересень–жовтень 2025 р. | Виконано |
| 3     | Розгляд прикладів обчислення натуральної щільності множин та базових властивостей випадкових сум | Жовтень–листопад 2025 р. | Виконано |
| 4     | Дослідження асимптотичних оцінок різниці між повною та прорідженою сумами                        | Листопад–грудень 2025 р. | Виконано |
| 5     | Встановлення умов збіжності проріджених сум до нормального розподілу                             | Січень–лютий 2026 р.     | Виконано |
| 6     | Дослідження швидкості збіжності до нормального розподілу (оцінки типу нерівності Беррі–Ессеєна)  | Лютий–березень 2026 р.   | Виконано |
| 7     | Проведення чисельних експериментів та побудова графіків отриманих результатів                    | Березень–квітень 2026 р. | Виконано |
| 8     | Підготовка доповіді та апробація результатів роботи на наукових конференціях                     | Квітень–травень 2026 р.  | Виконано |
| 9     | Завершальне редагування бакалаврської роботи та підготовка до захисту                            | Травень–червень 2026 р.  | Виконано |

Студентка

\_\_\_\_\_ Марія ДЕГТЯРЬОВА

Керівник

\_\_\_\_\_ Ірина НІЩЕНКО

## РЕФЕРАТ

Кваліфікаційна робота містить: 63 стор., 11 рисунків, 18 джерел.

У роботі досліджуються проріджені суми незалежних однаково розподілених випадкових величин, множина індексів яких є підмножиною множини натуральних чисел. Встановлено асимптотичну еквівалентність повної та прорідженої сум у просторі  $L^2$ . Отримано умови збіжності нормованих проріджених сум до нормального розподілу. Доведено нерівність Беррі–Ессеєна, що визначає оцінку швидкості збіжності проріджених сум до нормального розподілу. Наведено чисельні ілюстрації отриманих результатів.

ПРОРІДЖЕНІ СУМИ, ЦЕНТРАЛЬНА ГРАНИЧНА ТЕОРЕМА,  
НОРМАЛЬНИЙ РОЗПОДІЛ, НЕРІВНІСТЬ БЕРРІ–ЕССЕЕНА,  
АСИМПТОТИЧНА ЕКВІВАЛЕНТНІСТЬ, ВИПАДКОВІ СУМИ

## ABSTRACT

The qualification thesis consists of 63 pages, 11 figures, and 18 references.

The work investigates thinned sums of independent identically distributed random variables, where the index set is a subset of the natural numbers. Asymptotic equivalence between the full and thinned sums in the  $L^2$ . Conditions for convergence of normalized thinned sums to the normal distribution are obtained. A Berry–Esseen inequality is proved, providing an estimate of the rate of convergence of thinned sums to the normal distribution. Numerical illustrations of the obtained results are presented.

THINNED SUMS, CENTRAL LIMIT THEOREM, NORMAL DISTRIBUTION, BERRY–ESSEEN INEQUALITY, ASYMPTOTIC EQUIVALENCE, RANDOM SUMS

## ЗМІСТ

|                                                                                        |    |
|----------------------------------------------------------------------------------------|----|
| Перелік умовних позначень, скорочень і термінів .....                                  | 8  |
| Вступ.....                                                                             | 9  |
| 1 Огляд літератури .....                                                               | 11 |
| 1.1 Історична довідка .....                                                            | 11 |
| 1.2 Актуальність дослідження .....                                                     | 13 |
| Висновки до розділу 1 .....                                                            | 13 |
| 2 Натуральна щільність множин, перестановки та проріджені випадкові суми .....         | 14 |
| 2.1 Поняття натуральної щільності .....                                                | 14 |
| 2.2 Приклади обчислення натуральної щільності .....                                    | 15 |
| 2.3 Частотна характеристика перестановок .....                                         | 22 |
| 2.4 Поняття повної та прорідженої випадкової суми .....                                | 23 |
| 2.5 Ілюстрація впливу натуральної щільності множини на поведінку проріджених сум ..... | 24 |
| Висновки до розділу 2 .....                                                            | 26 |
| 3 Асимптотичні властивості проріджених випадкових сум .....                            | 28 |
| 3.1 Центральна гранична теорема для проріджених сум .....                              | 28 |
| 3.2 Асимптотика кількості вилучених індексів .....                                     | 38 |
| 3.3 Асимптотична оцінка різниці повної та прорідженої сум .....                        | 41 |
| 3.4 Оцінка швидкості збіжності: нерівність Беррі–Ессеєна .....                         | 43 |
| 3.5 Швидкість збіжності проріджених сум до нормального розподілу .....                 | 44 |
| 3.6 Оцінювання параметра прорідження $\beta$ .....                                     | 47 |
| 3.7 Результати моделювання .....                                                       | 49 |
| Висновки до розділу 3 .....                                                            | 59 |
| Висновки .....                                                                         | 61 |
| Перелік посилань .....                                                                 | 62 |

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ЦГТ - центральна гранична теорема

$\mathbb{N}$  — множина натуральних чисел

$\emptyset$  — порожня множина

$\sim$  — асимптотична еквівалентність ( $a_n \sim b_n$  тобто  $\frac{a_n}{b_n} \rightarrow 1$  при  $n \rightarrow \infty$ )

н.о.р. — незалежно однаково розподілені

в.в. — випадкові величини

## ВСТУП

**Актуальність дослідження.** Дослідження асимптотичних властивостей випадкових сум є одним із центральних напрямів теорії ймовірностей. У класичному випадку добре відомі результати щодо збіжності нормованих сум незалежних однаково розподілених випадкових величин до нормального розподілу. Водночас у багатьох прикладних задачах розглядаються не повні суми, а їх проріджені аналоги, тобто суми за підмножинами індексів. Такі ситуації природно виникають у задачах із пропущеними даними, нерегулярною вибіркою та вибіркоvim спостереженням. Вплив структури множини індексів, зокрема її натуральної щільності, на поведінку відповідних сум досліджено недостатньо повно, що зумовлює актуальність даної роботи.

**Метою дослідження** є порівняння повної та прорідженої сум, встановлення умов їх асимптотичної еквівалентності, а також дослідження граничного розподілу проріджених сум. Окрему увагу приділено оцінці швидкості збіжності до нормального розподілу.

### **Задачі дослідження:**

- 1) розглянути приклади обчислення натуральної щільності множин;
- 2) дослідити властивості повних та проріджених випадкових сум;
- 3) отримати асимптотичні оцінки різниці між повною та прорідженою сумами;
- 4) встановити умови збіжності проріджених сум до нормального розподілу;
- 5) дослідити швидкість збіжності до нормального розподілу та застосувати оцінки типу нерівності Беррі–Ессеєна;
- 6) проаналізувати вплив параметра, що визначає асимптотичну поведінку множини індексів, на поведінку сум;
- 7) провести чисельні експерименти та візуалізацію отриманих результатів.

*Об'єктом дослідження* є випадкові проріджені суми незалежних випадкових величин.

*Предметом дослідження* є асимптотичні властивості проріджених випадкових сум залежно від структури множини індексів.

*Методи дослідження* включають методи теорії ймовірностей, зокрема граничні теореми, оцінки швидкості збіжності, а також методи математичної статистики та чисельного моделювання.

**Наукова новизна** полягає у дослідженні впливу натуральної щільності та структури індексної множини на асимптотичну поведінку проріджених випадкових сум, а також у встановленні умов збіжності та оцінок швидкості збіжності до нормального розподілу.

**Практичне значення** одержаних результатів полягає у можливості застосування отриманих оцінок та підходів у задачах обробки даних із пропусками, моделюванні випадкових процесів та фінансовій математиці, де розглядаються нерегулярні або вибірккові спостереження.

**Апробація результатів та публікації.** Основні результати роботи доповідалися на семінарі для студентів з додаткових глав теорії ймовірностей та функціонального аналізу при відділі Інституту математики НАН України та XXIV Всеукраїнській науково-практичній конференції студентів, аспірантів та молодих вчених “Теоретичні і прикладні проблеми фізики, математики та інформатики”.

## 1 ОГЛЯД ЛІТЕРАТУРИ

У даному наведено короткий історичний огляд розвитку теорії випадкових сум та граничних теорем, а також окреслено основні підходи до аналізу асимптотичної поведінки сум випадкових величин з урахуванням структури множини індексів.

### 1.1 Історична довідка

Історія дослідження асимптотичних властивостей випадкових сум тісно пов'язана з граничними теоремами теорії ймовірностей та аналітичної теорії чисел.

Поняття натуральної щільності множин виникло в аналітичній теорії чисел у зв'язку з дослідженням асимптотичної поведінки підмножин натуральних чисел. У 2009 році у роботі Грекоса [1] було систематизовано низку відкритих проблем, пов'язаних із поняттям щільності множин натуральних чисел. Зокрема, було сформульовано ряд питань щодо властивостей натуральної щільності та її узагальнень, що відображає активний розвиток цієї теорії в сучасній математиці.

Перші результати центральної граничної теореми, беруть початок у роботах Лапласа, який використовував їх як інструмент для дослідження інших задач теорії ймовірностей. Пуассон доповнив цей напрям, розглядаючи контрприклад та вводячи ранні уявлення про випадкові величини. Подальший розвиток пов'язаний із Діріхле, Коші та Чебишовим, які досліджували похибки вимірювань і перші форми граничних результатів. Марков і Ляпунов сформулювали та довели класичні версії ЦГТ, зокрема у формі Ляпунова. У 1920 році Поліа ввів сам термін «центральна гранична теорема». Ліндеберг узагальнив умови збіжності, Бернштейн сформулював важливу допоміжну лему, яка

встановлює, що при достатньо слабкій залежності між випадковими величинами їхня сума асимптотично поводить ся як сума майже незалежних величин, що дозволяє застосовувати центральну граничну теорему. Згодом Леві та Феллер встановили необхідні й достатні умови ЦГТ, розвиваючи теорію різними методами, що остаточно закріпило її фундаментальний статус у теорії ймовірностей [2].

Подальший розвиток цієї теорії був пов'язаний із дослідженням точності наближення розподілів до нормального

Важливим етапом стало дослідження швидкості збіжності до нормального розподілу. Перші кількісні оцінки швидкості збіжності були отримані Ляпуновим (1900–1901). У 1940-х роках з'явилася класична нерівність Беррі–Ессеєна, яка дає універсальну та оптимальну за порядком оцінку збіжності [3].

У подальших роботах значна увага приділялася уточненню абсолютної константи в нерівності Беррі–Ессеєна. Серед сучасних результатів важливе місце займають роботи Шевцової [4], у яких отримано одні з найточніших оцінок константи для класичної центральної граничної теореми.

Окремим напрямом узагальнення класичних моделей є врахування неповноти або структури вибірки доданків, що приводить до поняття проріджених випадкових сум.

Ідея прорідження випадкових сум полягає у розгляді сум лише частини елементів послідовності, що визначається певним правилом вибору індексів. У ймовірнісному контексті це відповідає вилученню частини доданків або вибору підпослідовності індексів, що призводить до проріджених випадкових сум. Такі конструкції є узагальненням класичних випадкових сум і використовуються для дослідження впливу структури множини індексів на асимптотичну поведінку сум.

Отримані результати слугують основою для формулювання подальших задач дослідження.

## 1.2 Актуальність дослідження

Актуальність даної роботи зумовлена необхідністю:

- 1) уточнення умов, за яких проріджені суми зберігають класичні граничні властивості (зокрема збіжність до нормального розподілу);
- 2) дослідження швидкості збіжності та точності нормальних апроксимацій;
- 3) аналізу впливу параметрів, що задають структуру множини індексів, на асимптотичну поведінку проріджених сум;
- 4) застосування отриманих результатів у статистиці, моделюванні випадкових процесів та фінансовій математиці, де природно виникають ситуації з “неповними” або вибілковими сумами.

Таким чином, дослідження проріджених випадкових сум є актуальним і перспективним напрямом сучасної теорії ймовірностей, який поєднує класичні граничні теореми з новими підходами до аналізу структурованих підмножин індексів.

## Висновки до розділу 1

У даному розділі було розглянуто основні етапи розвитку теорії випадкових сум та їх граничних властивостей. Проведено аналіз класичних результатів, пов’язаних із центральною граничною теоремою та оцінками типу Беррі–Ессеєна, а також висвітлено роль натуральної щільності множин у дослідженні підмножин натуральних чисел.

Показано, що поєднання цих підходів приводить до задач про проріджені випадкові суми, для яких поведінка може істотно відрізнятися від класичного випадку. Обґрунтовано актуальність подальшого дослідження, зокрема в напрямі вивчення впливу структури індексної множини та її щільності на асимптотичні властивості сум, а також оцінювання швидкості їх збіжності.

## 2 НАТУРАЛЬНА ЩІЛЬНІСТЬ МНОЖИН, ПЕРЕСТАНОВКИ ТА ПРОРІДЖЕНІ ВИПАДКОВІ СУМИ

У цьому розділі розглядається поняття натуральної щільності множин та приклади її обчислення. Вводяться повні та проріджені випадкові суми, а також досліджується їх поведінка при різних типах прорідження множини індексів. Завершується розділ чисельними прикладами, що ілюструють залежність між натуральною щільністю множини та відхиленням проріджених сум від повних.

### 2.1 Поняття натуральної щільності

Наведемо означення натуральної щільності множини, яке дозволяє описувати середню частку елементів у підмножинах натуральних чисел.

**Означення 2.1** (Щільність множини [5]). Нехай  $A \subseteq \mathbb{N}$  є довільною підмножиною множини натуральних чисел і нехай  $(\lambda_n)_{n \geq 1}$  є довільною послідовністю невід'ємних чисел таких, що  $\sum_{n=1}^{\infty} \lambda_n = \infty$ . Визначимо

$$d(A; x) = \frac{\sum_{\substack{a \leq x \\ a \in A}} \lambda_a}{\sum_{n \leq x} \lambda_n}.$$

Якщо існує границя

$$d(A) = \lim_{x \rightarrow \infty} d(A; x),$$

то число  $d(A)$  називається *щільністю множини  $A$* .

**Означення 2.2** (Натуральна щільність множини [5]). Нехай  $A \subseteq \mathbb{N}$ . Натуральна щільність множини  $A$  визначається як

$$d(A) = \lim_{n \rightarrow \infty} d_n(A),$$

де  $d_n(A) = \frac{1}{n} |A \cap \{1, \dots, n\}|$ , якщо ця границя існує.

## 2.2 Приклади обчислення натуральної щільності

**Означення 2.3** (Перестановка [7]). Нехай  $X$  є скінченною або зліченною множиною. *Перестановкою* множини  $X$  називають бієкцію  $\sigma : X \rightarrow X$ , тобто відображення, яке кожному елементу  $x \in X$  ставить у відповідність єдиний елемент  $\sigma(x) \in X$  так, що з  $x_1 \neq x_2$  випливає  $\sigma(x_1) \neq \sigma(x_2)$

Для перестановки  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$  на  $\{1, 2, \dots, n\}$  позначимо через

$$K_n(\sigma) = \sum_{k,l=1}^n \mathbf{1}\{k = \sigma(l)\}$$

потужність перетину множини  $\{1, \dots, n\}$  та її образу при відображенні  $\sigma$  і покладемо

$$K(\sigma) = \lim_{n \rightarrow \infty} \frac{1}{n} K_n(\sigma)$$

Розглянемо також натуральну щільність  $d(A)$  множини

$$A = \{k \in \mathbb{N} : \sigma(k) = k\}$$

нерухомих точок відображення  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ , яка показує, яку частку індексів перестановка залишає на місці. Це дозволяє оцінити, наскільки сильно перестановка порушує порядок елементів.

Нижче детально розглянемо обидві характеристики для різних типів перестановок.

### Тотожна перестановка.

Нехай  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$  є тотожною перестановкою, тобто  $\sigma(i) = i$  для довільного  $i \in \mathbb{N}$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n & \dots \\ 1 & 2 & 3 & \dots & n & \dots \end{pmatrix}$$

Оскільки для кожного  $\ell \in \mathbb{N}$  маємо  $\sigma(\ell) = \ell$ , то  $K_n(\sigma) = \sum_{k,\ell=1}^n \mathbf{1}\{k = \sigma(\ell)\} = n$ , а отже  $K(\sigma) = 1$ .

Множина нерухомих точок співпадає з  $\mathbb{N}$ :  $A = \{k \in \mathbb{N} : \sigma(k) = k\} = \mathbb{N}$ , а тому  $d(A) = 1$ .

Тотожна перестановка не змінює порядок індексів: обидві характеристики дають значення 1.

### Скінченна перестановка.

Розглянемо перестановку  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & r & r+1 & r+2 & \dots & n & \dots \\ \sigma(1) & \sigma(2) & \dots & \sigma(r) & r+1 & r+2 & \dots & n & \dots \end{pmatrix}$$

Переставлено лише перші  $r$  елементів,  $\sigma(k) = k$  для всіх  $k > r$ . Оскільки для довільного  $n > r$

$$K_n(\sigma) = \sum_{k,\ell=1}^n \mathbf{1}\{k = \sigma(\ell)\} = n,$$

то  $K(\sigma) = 1$ .

З того, що для довільного  $n > r$  виконуються нерівності

$$n - r \leq |A \cap \{1, \dots, n\}| \leq n$$

випливає, що  $d(A) = 1$ .

Хоч порядок порушено на скінченній кількості індексів, їхній “внесок” зникає в границі, тому обидві характеристики дорівнюють 1.

## Перестановка пар.

Розглянемо перестановку  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$

$$\sigma(2k+1) = 2k+2, k = 0, 1, 2, \dots$$

$$\sigma(2k) = 2k-1, k = 1, 2, \dots$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \dots \\ 2 & 1 & 4 & 3 & 6 & 5 & 8 & 7 & \dots \end{pmatrix}$$

Обчислимо  $K(\sigma)$ . Для кожного  $\ell$  існує один індекс  $k$  такий, що  $k = \sigma(\ell)$ . Тому

$$K_n(\sigma) = \sum_{k, \ell=1}^n \mathbf{1}\{k = \sigma(\ell)\} = \begin{cases} n, & n \text{ парне,} \\ n-1, & n \text{ непарне,} \end{cases}$$

а отже  $K(\sigma) = 1$ .

Оскільки жоден елемент не є нерухомою точкою відображення  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ , то  $A = \emptyset$ , і  $d(A) = 0$ .

Цей приклад показує принципову різницю між величинами  $K(\sigma)$  і  $d(A)$ :

– Рівність  $K(\sigma) = 1$  не пов'язана зі збереженням порядку. Вона відображає лише той факт, що  $\sigma$  є бієкцією.

– Множина  $A$  є порожньою, тому її натуральна щільність дорівнює нулю  $d(A) = 0$ , тобто перестановка повністю руйнує локальний порядок натуральних чисел.

## Перестановка блоків зростаючої довжини $2^m$ .

Розглянемо перестановку  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ , яка працює з блоками зростаючої довжини  $2^m$ . Розіб'ємо множину  $\mathbb{N}$  на послідовні блоки довжини  $2^m$ . На кожному такому блоці перестановка  $\sigma$  здійснює циклічний зсув елементів на  $2^{m-1}$  позицій.

Перестановка виглядає так:

$$\sigma = \left( \begin{array}{cc|cccc|cccccccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & \dots \\ 2 & 1 & 5 & 6 & 3 & 4 & 11 & 12 & 13 & 14 & 7 & 8 & 9 & 10 & \dots \end{array} \right)$$

Позначимо через

$$N_k = 2 \cdot \sum_{m=0}^k 2^m = 2 \cdot (2^{k+1} - 1)$$

розділювачі між парами блоків. Наприклад,

$$N_0 = 2, \quad N_1 = 6.$$

Розглянемо спочатку повний блок довжини  $2^k$ . Для  $n = N_k$  беремо всі елементи до кінця  $k$ -го блоку. Тоді  $\sigma(\{1, \dots, N_k\}) = \{1, \dots, N_k\}$ , тобто:

$$K_{N_k}(\sigma) = N_k,$$

та

$$\lim_{k \rightarrow \infty} \frac{K_{N_k}(\sigma)}{N_k} = 1.$$

Розглянемо частковий блок. Якщо  $n = N_k + 2^{k+1}$  — середина  $k$ -го блоку, то

$$K_{N_k+2^{k+1}}(\sigma) = N_k,$$

і тоді

$$\lim_{k \rightarrow \infty} \frac{N_k}{N_k + 2^{k+1}} = \lim_{k \rightarrow \infty} \frac{2 \cdot (2^{k+1} - 1)}{3 \cdot 2^{k+1}} = \frac{2}{3}.$$

Отже, існують дві підпослідовності  $n$  з різними граничними значеннями: 1 та  $2/3$ . Тому границі  $K(\sigma)$  при  $n \rightarrow \infty$  не існує.

У даній перестановці нерухомих точок немає:

$$A = \emptyset, \quad d(A) = 0.$$

Цей приклад демонструє, що границя  $K(\sigma)$  при  $n \rightarrow \infty$  може не існувати, у той час як натуральна щільність множини нерухомих точок існує.

### Блокова перестановка з фіксованими елементами.

**Розглянемо приклад:**

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & \dots \\ 1 & 2 & 13 & 4 & 5 & 16 & 7 & 8 & 19 & 10 & 11 & 22 & 13 & 14 & 25 & 16 & \dots \end{pmatrix}$$

У наведеному записі видно ідею: ми намагаємося «зрушувати» кожне третє число на константу (тут додається 10). Однак такий підхід не працює для побудови перестановки (бієкції) на  $\mathbb{N}$ : наприклад, якщо покласти  $\sigma(3) = 13$  та одночасно  $\sigma(13) = 13$ , то отримаємо  $\sigma(3) = \sigma(13)$  — отже  $\sigma$  не є ін'єкцією.

Помітимо, що:

- Якщо «додавати 10» до кожного третього числа, то образи деяких чисел можуть співпасти з образом інших — це порушує ін'єкцію.
- Щоб уникнути прямих зіткнень, можна додавати число, що кратне 3. Така операція зберігає класи за модулем 3 (тобто образ кожного числа, що кратне 3, теж буде кратний 3).

### Приклад ін'єкції, яка не є сюр'єкцією:

$$\sigma(3k-2) = 3k-2, \quad \sigma(3k-1) = 3k-1, \quad \sigma(3k) = 3k+3 \quad (k = 1, 2, \dots)$$

Тут перші два елементи в кожній трійці залишаються фіксованими, а кожний третій «зсуваємо» на наступний третій. Це відображення ін'єктивне (образи різні), але не є сюр'єкцією, бо, наприклад, число 3 не лежить в образі (нічого не відображається на 3), якщо ми починаємо з  $3 \mapsto 6, 6 \mapsto 9$  тощо.

**Розглянемо ще приклад.**

Розділимо  $\mathbb{N}$  на блоки по три:

$$B_n = \{3n - 2, 3n - 1, 3n\}, \quad n = 1, 2, \dots$$

Визначимо перестановку  $\sigma$  таким чином:

$$\sigma(3n - 2) = 3n - 2, \quad \sigma(3n - 1) = 3n - 1,$$

а для третіх елементів зробимо попарну перестановку:

$$\sigma(3(2m - 1)) = 3(2m), \quad \sigma(3(2m)) = 3(2m - 1) \quad (m = 1, 2, \dots).$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & \dots \\ 1 & 2 & 6 & 4 & 5 & 3 & 7 & 8 & 12 & 10 & 11 & 9 & 13 & 14 & \dots \end{pmatrix}$$

Іншими словами, третій елемент першого блоку міняється місцями з третім елементом другого блоку, третій елемент третього — з четвертим і т.д.

Ця  $\sigma$  також не є прикладом бієкції з границею відмінною від 1, тобто  $K(\sigma) = 1$ . Також знайдемо натуральну щільність

$$|A \cap \{1, \dots, 3n\}| = 2n, \quad d_{3n}(A) = \frac{|A \cap \{1, \dots, 3n\}|}{3n} = \frac{2n}{3n} = \frac{2}{3}.$$

**Перестановка блоків лінійної довжини  $n$ .**

Розглянемо перестановку, яка діє на блоках довжини

$$1, 2, 3, \dots, k.$$

Позначимо загальну довжину перших  $k$  блоків:

$$N_k = \sum_{m=1}^k 2m = 2 \cdot \frac{k(k+1)}{2} = k(k+1).$$

Спочатку розглянемо кінці повних пар блоків. Якщо  $n = N_k$ , тобто розглядається кінець повного блоку, то всі переставлені елементи залишаються в межах  $1, \dots, n$ . Тоді

$$K_n(\sigma) = N_k,$$

і отже

$$\lim_{k \rightarrow \infty} \frac{K_n(\sigma)}{N_k} = \lim_{k \rightarrow \infty} \frac{N_k}{N_k} = 1.$$

Тепер розглянемо точку всередині наступного блоку. Розглянемо  $n = N_k + (k + 1)$ , тобто ми перебуваємо всередині наступного блоку довжини  $k + 1$ . На цьому відрізку частина елементів уже виходить за межі попередніх блоків, але кількість “правильних” співпадінь лишається  $N_k$ . Тоді

$$\frac{K_n(\sigma)}{n} = \frac{N_k}{N_k + (k + 1)} = \frac{k(k + 1)}{k(k + 1) + (k + 1)} = \frac{k^2 + k}{k^2 + k + 2} \xrightarrow{k \rightarrow \infty} 1.$$

Звідси можна зробити висновок, що у випадку лінійно зростаючих блоків обидві підпоследовності ( $n = N_k$  та  $n = N_k + (k + 1)$ ) дають однакову границю 1. Отже,

$$\lim_{n \rightarrow \infty} \frac{K_n(\sigma)}{n} = 1.$$

Натуральна щільність нерухомих точок у цій перестановці дорівнює нулю, оскільки жоден елемент не залишається на місці.

### **Перестановка блоків степеневі довжини $\lfloor n^\alpha \rfloor$ .**

Нехай довжина  $m$ -го блоку дорівнює  $m^\alpha$ , де  $\alpha > 0$  — фіксоване число. Тоді сумарна довжина перших  $k$  блоків має вигляд

$$N_k = \sum_{m=1}^k m^\alpha = D_0 k^{\alpha+1} + D_1 k^\alpha + \dots + D_\alpha k.$$

Знайдемо границю для  $K_n(\sigma)$ . Зауважимо, що

$$\frac{K_n(\sigma)}{N_k + (k+1)^\alpha} = \frac{N_k}{N_k + (k+1)^\alpha} = \frac{1}{1 + \frac{(k+1)^\alpha}{N_k}}.$$

Оскільки

$$\frac{(k+1)^\alpha}{N_k} \sim \frac{k^\alpha}{D_0 k^{\alpha+1}} = \frac{1}{D_0 k} \xrightarrow{k \rightarrow \infty} 0.$$

то

$$\frac{N_k}{N_k + (k+1)^\alpha} \xrightarrow{k \rightarrow \infty} 1.$$

Отже, як і у випадку лінійних блоків, для будь-якої степеневі швидкості росту блоків  $m^\alpha$  маємо

$$\lim_{n \rightarrow \infty} \frac{K_n(\sigma)}{n} = 1.$$

Фіксованих точок знову немає, тому

$$d(A) = 0.$$

Таким чином, навіть для значно нерівномірних, швидко зростаючих блоків перестановка повністю “плутає” індекси (щільність множини нерухомих точок нульова), але характеристика  $K_n(\sigma)$  поводить себе стабільно, і при нормуванні завжди прямує до 1.

**Зауваження.** Надалі вважатимемо, що  $A \subseteq \mathbb{N}$  є такою підмножиною натуральних чисел, для якої  $d(A) = 1$ .

## 2.3 Частотна характеристика перестановок

**Теорема 2.1.** Нехай для деякого  $\alpha < 1$

$$\limsup_{j \rightarrow \infty} \frac{|j - \sigma(j)|}{j^\alpha} < \infty. \quad (2.1)$$

Тоді існує границя

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbf{1}\{k = \sigma(l)\}.$$

**Доведення.** З умови (2.1) випливає, що існує константа  $C > 0$  така, що для довільного  $j > 1$  виконується  $|j - \sigma(j)| \leq Cj^\alpha$ . Тоді

$$K_n(\sigma) = \frac{1}{n} \sum_{k,l=1}^n \mathbf{1}\{k = \sigma(l)\} \geq \frac{1}{n}(n - Cn^\alpha)$$

З нерівностей

$$\frac{1}{n}(n - Cn^\alpha) \leq K_n(\sigma) \leq 1$$

випливає, що

$$K_n(\sigma) \rightarrow 1, n \rightarrow \infty$$

□

## 2.4 Поняття повної та прорідженої випадкової суми

Нехай  $\{X_i\}_{i \geq 1}$  є послідовністю незалежних однаково розподілених випадкових величин.

**Означення 2.4.** Повною сумою називатимемо суму вигляду

$$S_n = \sum_{i=1}^n X_i.$$

**Означення 2.5.** Нехай  $A \subset \mathbb{N}$  — довільна підмножина натуральних чисел. Прорідженою сумою, що відповідає множині  $A$ , називається сума

$$S_n(A) = \sum_{i=1}^n \mathbf{1}_{\{i \in A\}} X_i.$$

Тобто у прорідженій сумі враховуються лише ті доданки, індекси яких належать множині  $A$ .

У випадку випадкової кількості доданків аналогічно розглядаються випадкові суми

$$S_Y = \sum_{i=1}^Y X_i, \quad S_Y(A) = \sum_{i=1}^Y \mathbb{1}_{\{i \in A\}} X_i,$$

де  $Y$  є невід’ємною цілочисельною випадковою величиною.

Позначимо через

$$R_Y = S_Y - S_Y(A) = \sum_{i=1}^Y \mathbb{1}_{\{i \notin A\}} X_i$$

різницю між повною та прорідженою сумами.

## 2.5 Ілюстрація впливу натуральної щільності множини на поведінку проріджених сум

У цьому розділі наведено графічну ілюстрацію поведінки траєкторій повної суми  $S_k$  та проріджених сум  $S_k(A)$  для незалежних випадкових величин  $X_i \mathcal{N}(0,1)$ .

Розглянуто три приклади множин  $A$ :

- 1) Видалено кожний 100-й елемент ( $d(A) = 0.99$ );
- 2) Видалено елементи, індекси яких є квадратами натуральних чисел ( $d(A) = 1$ );
- 3) Залишено кожний другий елемент ( $d(A) = 0.5$ ).

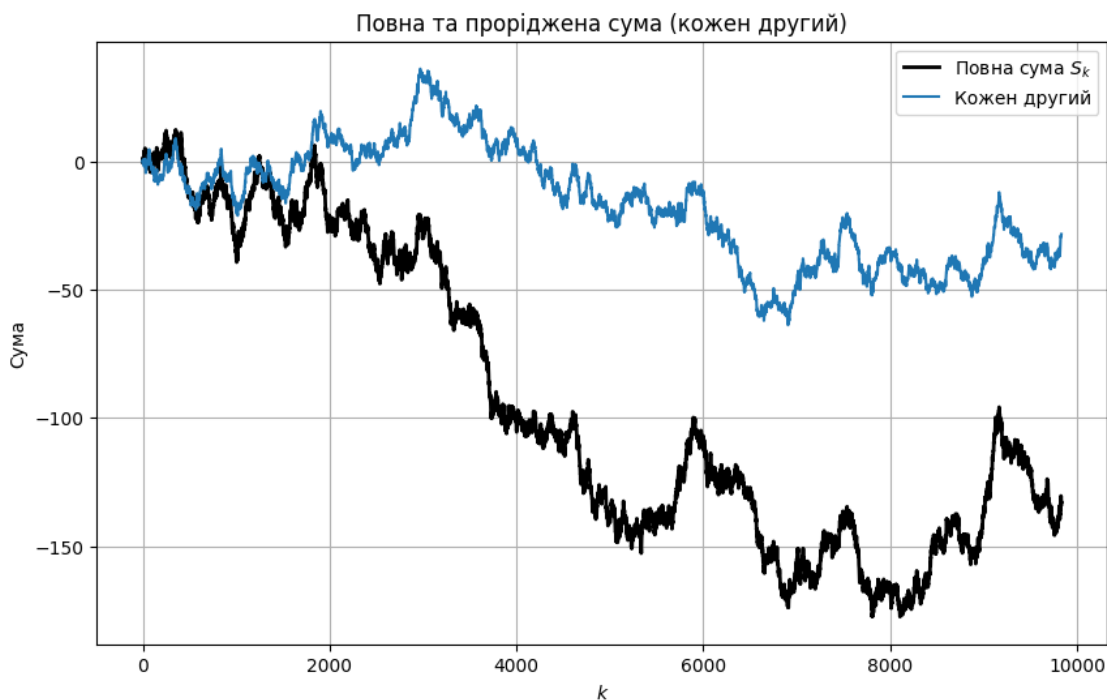
На кожному рисунку наведено траєкторії повної та прорідженої сум, що дозволяє оцінити вплив прорідження на поведінку часткових сум.



**Рисунок 2.1** – Повна сума  $S_k$  та проріджена сума з видаленням кожного 100-го елемента.



**Рисунок 2.2** – Повна сума  $S_k$  та проріджена сума з видаленням елементів з квадратними індексами.



**Рисунок 2.3** – Повна сума  $S_k$  та проріджена сума з кожним другим елементом.

Як видно на рисунках 2.1, 2.2, при щільності  $d(A)$ , близькій до 1, проріджена сума майже співпадає з повною сумою, тоді як на рис.2.3 при щільності  $d(A) = 0.5$  різниця стає добре помітною.

Ці спостереження свідчать про суттєвий вплив щільності множини  $A$  на поведінку проріджених сум.

Зокрема, виникає природне питання про те, наскільки близькими є повна та проріджена суми при  $d(A)$  близькій до 1, та як саме ця наближення проявляється в асимптотичному сенсі.

## Висновки до розділу 2

У даному розділі було введено поняття натуральної щільності множини та розглянуто її основні властивості на прикладах різних типів перестановок.

Зокрема, показано, як структура множини індексів впливає на її

щільність та як ця щільність змінюється при різних способах упорядкування натуральних чисел.

Окрему увагу приділено поняттю повних та проріджених випадкових сум, які виникають при виключенні частини доданків відповідно до заданої множини індексів.

Проведені чисельні експерименти демонструють, що поведінка проріджених сум суттєво залежить від щільності відповідної множини: при щільностях, близьких до одиниці, траєкторії сум майже не відрізняються, тоді як при зменшенні щільності спостерігаються виражені відхилення.

Таким чином, щільність множини виступає ключовим параметром, що визначає ступінь відхилення проріджених сум від повних.

### 3 АСИМПТОТИЧНІ ВЛАСТИВОСТІ ПРОРІДЖЕНИХ ВИПАДКОВИХ СУМ

У попередньому розділі було введено основні об'єкти дослідження — проріджені випадкові суми  $S_Y(A)$  з випадковою кількістю доданів  $Y \sim \text{Poisson}(\lambda)$ , а також поняття натуральної щільності множин, яке визначає структуру множини індексів  $A \subset \mathbb{N}$ .

У даному розділі досліджуються асимптотичні властивості цих об'єктів у граничному режимі  $\lambda \rightarrow \infty$ . Основна увага приділяється поведінці проріджених сум як випадкових процесів, їх нормуванню та граничним розподілам.

Зокрема, встановлюється центральна гранична теорема для проріджених сум, досліджується асимптотична еквівалентність повної та прорідженої сум, а також отримуються оцінки швидкості збіжності до нормального розподілу. Окремо розглядається задача оцінювання параметра прорідження на основі моментних характеристик спостережуваних сум.

У цьому розділі через  $|A_Y|$ ,  $|A_k|$  позначатимемо відповідно кількості елементів випадкової множини  $A \cap \{1, \dots, Y\}$ , та детермінованої множини  $A \cap \{1, \dots, k\}$ ,

#### 3.1 Центральна гранична теорема для проріджених сум

**Зауваження.** Надалі будемо вважати, що:

$\{X_i\}_{i \geq 1}$  — н.о.р. в.в;

$\mathbb{E}X_1 = 0$ ,  $\text{Var}(X_1) = 1$ ,  $\mathbb{E}|X_1|^3 < \infty$ ;

$Y \sim \text{Poisson}(\lambda)$ ;

$A \subset \mathbb{N}$  — фіксована множина з натуральною щільністю  $d(A)$ .

**Теорема 3.1.** Нехай  $A \subset \mathbb{N}$ . Тоді має місце збіжність за розподілом

$$\frac{S_Y(A)}{\sqrt{|A_Y|}} \xrightarrow{d} \mathcal{N}(0,1), \quad \lambda \rightarrow \infty.$$

**Доведення.** Розглянемо характеристичну функцію в.в  $\frac{S_Y(A)}{\sqrt{|A_Y|}}$

$$\varphi_{\frac{S_Y(A)}{\sqrt{|A_Y|}}}(t) = \mathbb{E} \left[ \exp \left( it \frac{S_Y(A)}{\sqrt{|A_Y|}} \right) \right].$$

Обчислимо спочатку умовне математичне сподівання  $\mathbb{E} \left( e^{it \frac{S_Y(A)}{\sqrt{|A_Y|}}} \mid Y = n \right)$ .

Маємо:

$$\begin{aligned} \mathbb{E} \left( e^{it \frac{S_Y(A)}{\sqrt{|A_Y|}}} \mid Y = n \right) &= \mathbb{E} \left( e^{\frac{it}{\sqrt{|A_n|}} \sum_{j=1}^n \mathbb{1}_{\{j \in A\}} X_j} \mid Y = n \right) \\ &= \prod_{j \in A} \varphi \left( \frac{t}{\sqrt{|A_n|}} \right) = \left( \varphi \left( \frac{t}{\sqrt{|A_n|}} \right) \right)^{|A_n|}, \end{aligned}$$

де  $\varphi(t)$  є характеристичною функцією випадкової величини  $X_1$

Врахувавши, що  $\mathbb{E}(X_1) = 0$ ,  $\mathbb{E}(X_1^2) = \text{Var}(X_1) = 1$ , можемо записати

$$\varphi \left( \frac{t}{\sqrt{|A_n|}} \right) = 1 - \frac{t^2}{2|A_n|} + o \left( \frac{1}{|A_n|} \right), \quad n \rightarrow \infty,$$

а тоді

$$\ln \varphi \left( \frac{t}{\sqrt{|A_n|}} \right) = -\frac{t^2}{2|A_n|} + o \left( \frac{1}{|A_n|} \right), \quad n \rightarrow \infty.$$

Звідси випливає, що

$$\left( \varphi \left( \frac{t}{\sqrt{|A_n|}} \right) \right)^{|A_n|} = e^{|A_n| \ln \varphi \left( \frac{t}{\sqrt{|A_n|}} \right)} \rightarrow e^{-\frac{t^2}{2}}, \quad n \rightarrow \infty.$$

За формулою повного математичного сподівання

$$\begin{aligned}\mathbb{E}\left(e^{it \frac{S_Y(A)}{\sqrt{|A_Y|}}}\right) &= \mathbb{E}\left(\mathbb{E}\left(e^{it \frac{S_Y(A)}{\sqrt{|A_Y|}}} \mid Y\right)\right) = \mathbb{E}\left[\left(\varphi\left(\frac{t}{\sqrt{|A_Y|}}\right)\right)^{|A_Y|}\right] \\ &= \sum_{k=0}^{\infty} P(Y = k) \left(\varphi\left(\frac{t}{\sqrt{|A_k|}}\right)\right)^{|A_k|}.\end{aligned}$$

Подамо праву частину у вигляді

$$\begin{aligned}\mathbb{E}\left(e^{it \frac{S_Y(A)}{\sqrt{|A_Y|}}}\right) &= e^{-t^2/2} + \sum_{k=0}^N P(Y = k) \left(\left(\varphi\left(\frac{t}{\sqrt{|A_k|}}\right)\right)^{|A_k|} - e^{-t^2/2}\right) \\ &\quad + \sum_{k=N+1}^{\infty} P(Y = k) \left(\left(\varphi\left(\frac{t}{\sqrt{|A_k|}}\right)\right)^{|A_k|} - e^{-t^2/2}\right).\end{aligned}\tag{3.1}$$

Для довільного  $\varepsilon > 0$  виберемо спочатку  $N^*$  так, щоб для всіх  $k > N^*$  виконувалася умова малості внутрішньої різниці:

$$\left|\left(\varphi\left(\frac{t}{\sqrt{|A_k|}}\right)\right)^{|A_k|} - e^{-t^2/2}\right| < \frac{\varepsilon}{2}.$$

Тоді третій доданок в (3.1) можна оцінити так:

$$\sum_{k=N^*+1}^{\infty} P(Y = k) \left|\left(\varphi\left(\frac{t}{\sqrt{|A_k|}}\right)\right)^{|A_k|} - e^{-t^2/2}\right| < \frac{\varepsilon}{2} \sum_{k=N^*+1}^{\infty} P(Y = k) \leq \frac{\varepsilon}{2}.$$

Для другого доданку в (3.1) маємо оцінку:

$$\sum_{k=0}^{N^*} P(Y = k) \left|\left(\varphi\left(\frac{t}{\sqrt{|A_k|}}\right)\right)^{|A_k|} - e^{-t^2/2}\right| \leq \sum_{k=0}^{N^*} P(Y = k) \cdot 2 = 2 P(Y \leq N^*).$$

Оскільки для фіксованого  $N^*$

$$P(Y \leq N^*) \rightarrow 0 \quad \text{при} \quad \lambda \rightarrow \infty,$$

то існує таке  $\lambda^*$ , що для всіх  $\lambda > \lambda^*$  виконується

$$2 P(Y \leq N^*) < \frac{\varepsilon}{2}.$$

Об'єднуючи обидві оцінки, отримуємо:

$$\left| \sum_{k=0}^{\infty} P(Y = k) \left( \left( \varphi \left( \frac{t}{\sqrt{|A_k|}} \right) \right)^{|A_k|} - e^{-t^2/2} \right) \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Тобто

$$\left| \mathbb{E} \left[ e^{itS_Y(A)/\sqrt{|A_Y|}} \right] - e^{-t^2/2} \right| < \varepsilon.$$

За теоремою непервності [12] зі збіжності характеристичних функцій випливає збіжність за розподілом

$$\frac{S_Y(A)}{\sqrt{|A_Y|}} \xrightarrow{d} \mathcal{N}(0,1), \quad \lambda \rightarrow \infty.$$

□

**Лема 3.1.**

$$\frac{\mathbb{E}(|A_Y|)}{\lambda} \rightarrow 1, \quad \lambda \rightarrow \infty.$$

**Доведення.** Маємо

$$\mathbb{E}[|A_Y|] = \sum_{k=0}^{\infty} |A_k| P(Y = k),$$

Оскільки  $\mathbb{E}[Y] = \sum_{k=0}^{\infty} k P(Y = k) = \lambda$ , то

$$\mathbb{E}[|A_Y|] = \lambda - \sum_{k=0}^{\infty} (k - |A_k|) P(Y = k).$$

і тоді

$$\frac{\mathbb{E}[|A_Y|]}{\lambda} = 1 - \frac{1}{\lambda} \sum_{k=0}^{\infty} (k - |A_k|) P(Y = k).$$

Розділимо суму в правій частині останньої рівності на дві частини

для деякого  $K > 0$ :

$$\sum_{k=0}^{\infty} (k - |A_k|) P(Y = k) = \sum_{k=0}^K (k - |A_k|) P(Y = k) + \sum_{k=K+1}^{\infty} k \left(1 - \frac{|A_k|}{k}\right) P(Y = k).$$

Оскільки

$$1 - \frac{|A_k|}{k} \rightarrow 0, k \rightarrow \infty,$$

то

$$\forall \varepsilon > 0, \exists K^* \forall k > K^* \quad \left|1 - \frac{|A_k|}{k}\right| \leq \frac{\varepsilon}{2}$$

і тому

$$\frac{1}{\lambda} \left| \sum_{k=K^*+1}^{\infty} k \left(1 - \frac{|A_k|}{k}\right) P(Y = k) \right| \leq \frac{\varepsilon}{2\lambda} \left| \sum_{k=K^*+1}^{\infty} k P(Y = k) \right| \leq \frac{\varepsilon}{2}.$$

Перший доданок:  $\sum_{k=0}^{K^*} (k - |A_k|) P(Y = k)$  обмежений константою  $C$  і при діленні на  $\lambda$  прямує до 0:

$$\frac{1}{\lambda} \sum_{k=0}^{K^*} (k - |A_k|) P(Y = k) \leq \frac{C}{\lambda} \rightarrow 0.$$

Об'єднавши обидві оцінки, отримуємо

$$\frac{1}{\lambda} \sum_{k=0}^{\infty} (k - |A_k|) P(Y = k) \rightarrow 0, \quad \lambda \rightarrow \infty.$$

Отже,

$$\frac{\mathbb{E}[|A_Y|]}{\lambda} \rightarrow 1, \quad \lambda \rightarrow \infty.$$

□

**Лема 3.2.**

$$\frac{\text{Var}(|A_Y|)}{\lambda} \rightarrow 1, \quad \lambda \rightarrow \infty.$$

**Доведення.**

За означенням дисперсії

$$\text{Var}(|A_Y|) = \mathbb{E}[|A_Y|^2] - (\mathbb{E}[|A_Y|])^2.$$

Для обчислення другого моменту

$$\mathbb{E}[|A_Y|^2] = \sum_{k=0}^{\infty} |A_k|^2 \mathbb{P}(Y = k),$$

скористаємося тотожністю

$$|A_k|^2 = k^2 - 2k(k - |A_k|) + (k - |A_k|)^2.$$

Тоді, користуючись лінійністю математичного сподівання, отримуємо

$$\mathbb{E}[|A_Y|^2] = \mathbb{E}[Y^2] - 2\mathbb{E}[Y(Y - |A_Y|)] + \mathbb{E}[(Y - |A_Y|)^2].$$

Оскільки  $Y \sim \text{Poisson}(\lambda)$ , то

$$\mathbb{E}[Y^2] = \lambda^2 + \lambda.$$

З попередньої теореми відомо, що

$$\mathbb{E}[|A_Y|] = \lambda + o(\lambda), \quad \lambda \rightarrow \infty.$$

Тобто існує така функція  $\varepsilon(\lambda)$ , що  $\frac{\varepsilon(\lambda)}{\lambda} \rightarrow 0$  і

$$\mathbb{E}[|A_Y|] = \lambda + \varepsilon(\lambda).$$

Тоді

$$(\mathbb{E}[|A_Y|])^2 = (\lambda + \varepsilon(\lambda))^2 = \lambda^2 + 2\lambda\varepsilon(\lambda) + \varepsilon(\lambda)^2.$$

З того, що

$$\frac{2\lambda\varepsilon(\lambda)}{\lambda^2} = 2\frac{\varepsilon(\lambda)}{\lambda} \rightarrow 0, \quad \frac{\varepsilon(\lambda)^2}{\lambda^2} = \left(\frac{\varepsilon(\lambda)}{\lambda}\right)^2 \rightarrow 0,$$

випливає

$$(\mathbb{E}[|A_Y|])^2 = \lambda^2 + o(\lambda^2).$$

Отже,

$$\text{Var}(|A_Y|) = \lambda - 2\mathbb{E}[Y(Y - |A_Y|)] + \mathbb{E}[(Y - |A_Y|)^2] + o(\lambda).$$

Оскільки згідно зі зробленими припущеннями  $\frac{|A_k|}{k} \rightarrow 1$  при  $k \rightarrow \infty$ , то  $k - |A_k| = o(k)$ . Для  $Y \sim \text{Poisson}(\lambda)$  це дає

$$\mathbb{E}[Y - |A_Y|] = o(\lambda), \quad \mathbb{E}[(Y - |A_Y|)^2] = o(\lambda^2).$$

Звідси, за нерівністю Коші–Буняковського,

$$\mathbb{E}[Y(Y - |A_Y|)] \leq \sqrt{\mathbb{E}[Y^2] \mathbb{E}[(Y - |A_Y|)^2]} = o(\lambda^2).$$

У підсумку

$$\text{Var}(|A_Y|) = \lambda + o(\lambda)$$

а отже,

$$\frac{\text{Var}(|A_Y|)}{\lambda} \rightarrow 1, \quad \lambda \rightarrow \infty,$$

що й стверджувалося. □

**Лема 3.3.** *Має місце збіжність за ймовірністю*

$$\frac{|A_Y|}{\lambda} \xrightarrow{P} 1, \quad \lambda \rightarrow \infty.$$

**Доведення.** Подамо різницю  $\frac{|A_Y|}{\lambda} - 1$  у вигляді

$$\frac{|A_Y|}{\lambda} - 1 = \left( \frac{|A_Y|}{\lambda} - \mathbb{E}\left[\frac{|A_Y|}{\lambda}\right] \right) + \left( \mathbb{E}\left[\frac{|A_Y|}{\lambda}\right] - 1 \right).$$

Перший доданок у дужках прямує до нуля за законом великих чисел у формі Чебишова [8]: для будь-якого  $\varepsilon > 0$  маємо

$$\Pr\left(\left|\frac{|A_Y|}{\lambda} - \mathbb{E}\left[\frac{|A_Y|}{\lambda}\right]\right| > \varepsilon\right) \leq \frac{\text{Var}\left(\frac{|A_Y|}{\lambda}\right)}{\varepsilon^2} = \frac{\frac{1}{\lambda^2} \text{Var}(|A_Y|)}{\varepsilon^2}.$$

Оскільки згідно з лемою 3.2  $\text{Var}(|A_Y|) = O(\lambda)$ , то

$$\frac{\text{Var}(|A_Y|)}{\lambda^2} \rightarrow 0, \quad \lambda \rightarrow \infty.$$

Отже,

$$\Pr\left(\left|\frac{|A_Y|}{\lambda} - \mathbb{E}\left[\frac{|A_Y|}{\lambda}\right]\right| > \varepsilon\right) \rightarrow 0, \quad \lambda \rightarrow \infty.$$

Другий доданок є детермінованим і, за лемою 3.1, теж прямує до 0:

$$\mathbb{E}\left[\frac{|A_Y|}{\lambda}\right] - 1 \rightarrow 0 \quad \lambda \rightarrow \infty.$$

Отже,

$$\frac{|A_Y|}{\lambda} - 1 \xrightarrow{P} 0,$$

звідси випливає твердження леми. □

**Лема 3.4.**

$$\mathbb{E}\left[\left(\frac{\lambda}{Y}\right)^{1/2}\right] \rightarrow 1, \quad \lambda \rightarrow \infty.$$

**Доведення.** Оскільки  $\mathbb{E}(Y) = \lambda$ ,  $\text{Var}(Y) = \lambda$ , то

$$\text{Var}\left(\frac{Y}{\lambda}\right) = \frac{1}{\lambda^2} \text{Var}(Y) = \frac{1}{\lambda} \rightarrow 0.$$

Тому за нерівністю Чебишова [8]

$$\frac{Y}{\lambda} \xrightarrow{P} 1, \quad \lambda \rightarrow \infty.$$

Функція

$$g(x) = x^{-1/2}$$

є неперервною в точці  $x = 1$ , тому за теоремою про неперервне відображення [13]

$$\left(\frac{Y}{\lambda}\right)^{-1/2} = \left(\frac{\lambda}{Y}\right)^{1/2} \xrightarrow{P} 1.$$

Позначимо

$$\xi_\lambda = \left(\frac{\lambda}{Y}\right)^{1/2}.$$

Візьмемо функцію

$$G(t) = t^3.$$

Тоді  $G$  є невід'ємною зростаючою функцією і

$$\frac{G(t)}{t} = t^2 \xrightarrow[t \rightarrow \infty]{} \infty.$$

Крім того,

$$\mathbb{E}G(\xi_\lambda) = \mathbb{E} \left[ \left(\frac{\lambda}{Y}\right)^{3/2} \right].$$

Покажемо, що

$$\sup_\lambda \mathbb{E} \left[ \left(\frac{\lambda}{Y}\right)^{3/2} \right] < \infty.$$

Оскільки

$$\frac{Y}{\lambda} \xrightarrow{P} 1,$$

то також

$$\left(\frac{\lambda}{Y}\right)^{3/2} \xrightarrow{P} 1.$$

Із того, що

$$\mathbb{E} \left( \frac{Y - \lambda}{\sqrt{\lambda}} \right)^2 = \frac{1}{\lambda} \mathbb{E}[(Y - \lambda)^2] = \frac{1}{\lambda} \text{Var}(Y) = \frac{1}{\lambda} \cdot \lambda = 1$$

впливає обмеженість відповідних моментів, тому

$$\sup_{\lambda} \mathbb{E}G(\xi_{\lambda}) < \infty.$$

Тоді за Лемою 3 [14, с. 190] сімейство

$$\left\{ \left( \frac{\lambda}{Y} \right)^{1/2} \right\}_{\lambda}$$

є рівномірно інтегровним.

Маємо

$$\left( \frac{\lambda}{Y} \right)^{1/2} \xrightarrow{P} 1,$$

причому випадкові величини є невід'ємними та інтегровними.

Отже, за Теоремою 5 [14, с. 189]

$$\mathbb{E} \left[ \left( \frac{\lambda}{Y} \right)^{1/2} \right] \rightarrow \mathbb{E}(1) = 1, \quad \lambda \rightarrow \infty.$$

□

**Теорема 3.2.**

$$\frac{S_Y(A)}{\sqrt{\lambda}} \xrightarrow{d} \mathcal{N}(0,1), \quad \lambda \rightarrow \infty.$$

**Доведення.** Оскільки за теоремою 3.1

$$\frac{S_Y(A)}{\sqrt{|A_Y|}} \xrightarrow{d} \mathcal{N}(0,1)$$

та за лемою 3.3

$$\sqrt{\frac{|A_Y|}{\lambda}} \xrightarrow{P} 1,$$

то за теоремою Слуцького [9]:

$$\frac{S_Y(A)}{\sqrt{\lambda}} = \frac{S_Y(A)}{\sqrt{|A_Y|}} \sqrt{\frac{|A_Y|}{\lambda}} \xrightarrow{d} \mathcal{N}(0,1).$$

□

### 3.2 Асимптотика кількості вилучених індексів

Для опису щільності прорідження введемо припущення на швидкість зростання кількості вилучених індексів.

**Теорема 3.3** (Оцінка щільності прорідження).

*Якщо*

$$k - |A_k| = O(k^\alpha), \quad k \rightarrow \infty,$$

*для деякого  $\alpha < 1$ , то*

$$\frac{\mathbb{E}[|A_Y|]}{\lambda} = 1 - O(\lambda^{\alpha-1}), \quad \lambda \rightarrow \infty.$$

**Доведення.** Маємо

$$1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} = \frac{1}{\lambda} \sum_{k=0}^{\infty} (k - |A_k|) \mathbb{P}(Y = k).$$

Використовуючи умову

$$k - |A_k| \leq Ck^\alpha, \quad k \geq 1. \quad (3.2)$$

отримуємо

$$1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} \leq \frac{C}{\lambda} \sum_{k=0}^{\infty} k^\alpha \mathbb{P}(Y = k) = \frac{C}{\lambda} \mathbb{E}[Y^\alpha].$$

Оскільки  $\alpha < 1$ , то за нерівністю Єнсена

$$\mathbb{E}[Y^\alpha] \leq (\mathbb{E}Y)^\alpha = \lambda^\alpha.$$

Отже,

$$1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} \leq C\lambda^{\alpha-1},$$

звідки випливає

$$1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} = O(\lambda^{\alpha-1}), \quad \lambda \rightarrow \infty.$$

Уточнимо отриману оцінку, встановивши точний результат.

**Теорема 3.4.** *Нехай виконується умова 3.2. Тоді*

$$\lambda^{1-\alpha} \left( 1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} \right) \longrightarrow 1, \quad \lambda \rightarrow \infty.$$

**Доведення.** Зі справедливості двосторонньої оцінки

$$\frac{1}{\lambda^\alpha} \mathbb{E}((Y-1)^\alpha) \leq \lambda^{1-\alpha} \left( 1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} \right) \leq \frac{1}{\lambda^\alpha} \mathbb{E}(Y^\alpha)$$

випливає, що достатньо довести, що

$$\frac{1}{\lambda^\alpha} \mathbb{E}(Y^\alpha) \rightarrow 1, \quad \frac{1}{\lambda^\alpha} \mathbb{E}((Y-1)^\alpha) \rightarrow 1, \quad \lambda \rightarrow \infty.$$

Позначимо  $X_\lambda = \frac{Y}{\lambda}$ ,  $Z_\lambda = \frac{Y-1}{\lambda}$ .

За центральної граничної теоремою

$$\frac{Y - \lambda}{\sqrt{\lambda}} \xrightarrow{d} N(0,1), \quad \lambda \rightarrow \infty.$$

Оскільки  $\frac{1}{\sqrt{\lambda}} \rightarrow 0, \lambda \rightarrow \infty$  то

$$X_\lambda = 1 + \frac{Y - \lambda}{\lambda} \xrightarrow{d} 1, \quad \lambda \rightarrow \infty.$$

Тоді

$$Z_\lambda = X_\lambda - \frac{1}{\lambda} \xrightarrow{d} 1, \quad \lambda \rightarrow \infty.$$

Оскільки  $X_\lambda \xrightarrow{d} 1$  та  $Z_\lambda \xrightarrow{d} 1$ , а функція  $f(x) = x^\alpha$  є неперервною, то за теоремою про неперервне відображення [13]

$$X_\lambda^\alpha \xrightarrow{d} 1, \quad Z_\lambda^\alpha \xrightarrow{d} 1, \quad \lambda \rightarrow \infty.$$

Застосуємо Лему 3 [14, с. 190] до послідовностей

$$\xi_\lambda = \left(X_\lambda^\alpha\right)_\lambda, \quad \eta_\lambda = \left(Z_\lambda^\alpha\right)_\lambda.$$

Візьмемо функцію

$$G(t) = t^{1/\alpha}.$$

Тоді  $G$  є невід'ємною зростаючою і, оскільки  $\alpha < 1$ ,

$$\frac{G(t)}{t} = t^{\frac{1}{\alpha}-1} \xrightarrow{t \rightarrow \infty} \infty.$$

Крім того,

$$\mathbb{E}G(\xi_\lambda) = \mathbb{E}(X_\lambda^\alpha)^{1/\alpha} = \mathbb{E}X_\lambda = \mathbb{E}\left(\frac{Y}{\lambda}\right) = 1,$$

тому

$$\sup_\lambda \mathbb{E}G(\xi_\lambda) < \infty.$$

Аналогічно,

$$\mathbb{E}G(\eta_\lambda) = \mathbb{E}Z_\lambda = \mathbb{E}\left(\frac{Y-1}{\lambda}\right) = 1 - \frac{1}{\lambda} \leq 1,$$

звідки

$$\sup_\lambda \mathbb{E}G(\eta_\lambda) < \infty.$$

Отже, сімейства  $\{X_\lambda^\alpha\}_\lambda, \{Z_\lambda^\alpha\}_\lambda$  є рівномірно інтегровними.

Маємо

$$X_\lambda^\alpha \xrightarrow{d} 1, \quad Z_\lambda^\alpha \xrightarrow{d} 1, \quad \lambda \rightarrow \infty,$$

причому ці випадкові величини невід'ємні і інтегровні.

Оскільки відповідні сімейства є рівномірно інтегровними, то за Теоремою 5 [14, с. 189]

$$\mathbb{E}(X_\lambda^\alpha) \rightarrow \mathbb{E}(1) = 1, \quad \mathbb{E}(Z_\lambda^\alpha) \rightarrow \mathbb{E}(1) = 1, \quad \lambda \rightarrow \infty.$$

Тобто

$$\frac{1}{\lambda^\alpha} \mathbb{E}(Y^\alpha) \rightarrow 1, \quad \lambda \rightarrow \infty, \quad \frac{1}{\lambda^\alpha} \mathbb{E}((Y-1)^\alpha) \rightarrow 1, \quad \lambda \rightarrow \infty.$$

Отже, за теоремою про стиснення [15]

$$\lambda^{1-\alpha} \left( 1 - \frac{\mathbb{E}[|A_Y|]}{\lambda} \right) \rightarrow 1, \quad \lambda \rightarrow \infty.$$

□

### 3.3 Асимптотична оцінка різниці повної та прорідженої сум

У цьому розділі покажемо, що проріджена сума  $S_Y(A)$  асимптотично еквівалентна повній сумі  $S_Y$  як у просторі  $L^2$ , так і за розподілом.

#### Оцінка в сенсі $L^2$ .

У цьому пункті дослідимо різницю між повною та прорідженою сумами в просторі  $L^2$ . Покажемо, що внесок від членів, які виключаються внаслідок прорідження, є асимптотично незначним порівняно з повною сумою.

$$\begin{aligned} \text{Var}(R_Y | Y) &= \text{Var}\left(\sum_{i=1}^Y \mathbf{1}_{\{i \notin A\}} X_i \mid Y\right) = \sum_{i=1}^Y \text{Var}(\mathbf{1}_{\{i \notin A\}} X_i | Y) \\ &= \sum_{i=1}^Y \mathbf{1}_{\{i \notin A\}}^2 \text{Var}(X_i) = \sum_{i=1}^Y \mathbf{1}_{\{i \notin A\}} = Y - |A_Y| \end{aligned}$$

Тоді за формулою повної дисперсії

$$\text{Var}(R_Y) = \text{Var}(\mathbb{E}[R_Y | Y]) + \mathbb{E}[\text{Var}(R_Y | Y)].$$

З того, що

$$\mathbb{E}[R_Y | Y] = \mathbb{E}\left[\sum_{i=1}^Y \mathbf{1}_{\{i \notin A\}} X_i \mid Y\right]$$

випливає, що

$$\text{Var}(\mathbb{E}[R_Y \mid Y]) = 0.$$

Тому

$$\text{Var}(R_Y) = \mathbb{E}[\text{Var}(R_Y \mid Y)] = \mathbb{E}\left[\sum_{i=1}^Y \mathbf{1}_{\{i \notin A\}}\right] = \mathbb{E}[Y - |A_Y|].$$

Згідно з лемою 3.1  $\mathbb{E}[Y - |A_Y|] = o(\lambda)$ ,  $\lambda \rightarrow \infty$ . Тоді

$$\|S_Y - S_Y(A)\|_{L^2} = \sqrt{\text{Var}(R_Y)} = o(\sqrt{\lambda}), \quad \lambda \rightarrow \infty.$$

Водночас,  $L^2$ -норма повної суми оцінюється так:

$$\|S_Y\|_{L^2} = \sqrt{\text{Var}(S_Y)} = \sqrt{\mathbb{E}[Y]} \sim \sqrt{\lambda}.$$

Отже,

$$\frac{\|S_Y - S_Y(A)\|_{L^2}}{\|S_Y\|_{L^2}} = \frac{\|R_Y\|_{L^2}}{\|S_Y\|_{L^2}} = \frac{o(\sqrt{\lambda})}{\sqrt{\lambda}} = o(1), \quad \lambda \rightarrow \infty,$$

### Порівняння за розподілом.

Позначимо

$$\tilde{S}_Y = \frac{S_Y}{\sqrt{\text{Var}(S_Y)}}, \quad \tilde{S}_Y(A) = \frac{S_Y(A)}{\sqrt{\text{Var}(S_Y(A))}}.$$

Оскільки  $\mathbb{E}X_i = 0$ , то  $\mathbb{E}(S_Y \mid Y) = 0$ ,  $\mathbb{E}(S_Y(A) \mid Y) = 0$ , і віднімати математичне сподівання не потрібно.

Крім того,  $\text{Var}(X_i) = 1$ , а також  $\text{Var}(S_Y) = \lambda$  і  $\text{Var}(S_Y(A)) = \lambda$  (лема 3.2), тому до кожної з цих сум можна застосувати ЦГТ. Тоді

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(\tilde{S}_Y \leq x) - \Phi(x)| = o(1), \quad \sup_{x \in \mathbb{R}} |\mathbb{P}(\tilde{S}_Y(A) \leq x) - \Phi(x)| = o(1).$$

Звідси, за нерівністю трикутника,

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(\tilde{S}_Y(A) \leq x) - \mathbb{P}(\tilde{S}_Y \leq x)| = o(1), \quad \lambda \rightarrow \infty.$$

Таким чином, проріджена сума  $S_Y(A)$  є асимптотично еквівалентною повній сумі  $S_Y$  у просторі  $L^2$  та за розподілом, а швидкість зближення визначається кількістю відсутніх індексів  $i \notin A$ .

### 3.4 Оцінка швидкості збіжності: нерівність Беррі–Ессеєна

ЦГТ описує ситуацію, коли нормована сума випадкових величин має граничний нормальний розподіл. Тоді виникає питання про те, з якою швидкістю розподіл нормалізованої суми наближається до стандартного нормального.

**Теорема 3.5** (Нерівність Беррі–Ессеєна). [3]

Нехай  $X_1, X_2, \dots, X_n$  — н.о.р.в.в. з математичним сподіванням  $\mu = \mathbb{E}X_i$ , дисперсією  $\sigma^2 = \mathbb{E}(X_i - \mu)^2 > 0$  та скінченним третім абсолютним моментом  $\beta_3 = \mathbb{E}|X_i - \mu|^3 < \infty$ .

Позначимо

$$S_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n (X_i - \mu),$$

а через  $F_n$  — функцію розподілу  $S_n$  та  $\Phi(x)$  — стандартна нормальна функція розподілу. Тоді існує універсальна константа  $C > 0$ , що не залежить від  $n$ , така що

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq C \frac{\beta_3}{\sigma^3 \sqrt{n}}.$$

### 3.5 Швидкість збіжності проріджених сум до нормального розподілу

У цьому розділі досліджується збіжність проріджених випадкових сум до нормального розподілу та оцінюється швидкість такої збіжності.

Наведена нижче теорема є одним із основних результатів роботи, оскільки встановлює оцінку типу Беррі–Ессеєна для проріджених випадкових сум і показує вплив параметра прорідження на точність нормального наближення.

**Теорема 3.6.** *Нехай  $(X_n)_{n \geq 1}$  є послідовністю н.о.р. випадкових величин, кожна з яких має рівномірний розподіл на інтервалі  $(-1,1)$ . Припустимо, що множина  $A \subseteq \mathbb{N}$  є такою, що  $1 - \frac{1}{n}|A \cap \{1, \dots, n\}| = O\left(\frac{1}{n^\alpha}\right)$ ,  $n \rightarrow \infty$  для деякого  $\alpha \in (0,1)$ . Тоді*

$$\sup_x \left| P\left(\frac{S_Y(A)}{\sqrt{\lambda}} \leq x\right) - \Phi_{0,1}(x) \right| \leq \frac{C_1}{\sqrt{\lambda^{\frac{1}{2}-\alpha}}} + \frac{C_2}{\sqrt{\lambda}},$$

де  $C_1, C_2$  є деякими додатними константами.

#### Доведення.

Зауважимо, що

$$\sum_{i=1}^n X_i = 2 \sum_{i=1}^n Z_i - n,$$

де  $Z_i = \frac{X_i+1}{2} \sim U(0,1)$ ,  $i = 1, \dots, n$ .

Тому

$$f_{\sum_{i=1}^k X_i}(x) = \frac{1}{2} f_{T_k}\left(\frac{x+k}{2}\right),$$

де  $T_k = \sum_{i=1}^k Z_i$ . Ця в.в має розподіл Ірвіна–Холла.

Оскільки

$$\mathbb{E}T_k = \frac{k}{2}, \quad \text{Var}(T_k) = \frac{k}{12},$$

то

$$\sup_x \left| \sqrt{\frac{k}{12}} f_{T_k}(x) - \varphi\left(\frac{x - k/2}{\sqrt{k/12}}\right) \right| \rightarrow 0,$$

де  $\varphi$  — щільність стандартного нормального розподілу.

Оскільки

$$\sup_x \varphi(x) = \frac{1}{\sqrt{2\pi}},$$

отримуємо

$$\sup_x f_{T_k}(x) \leq \frac{C}{\sqrt{k}}$$

де  $C > 0$  — деяка константа.

З того, що

$$f_{\sum_{i=1}^Y X_i/\sqrt{\lambda}}(t) = \sum_{k=1}^{\infty} P(Y = k) \frac{\sqrt{\lambda}}{2} f_{\sum_{i=1}^k Z_i}\left(\frac{\sqrt{\lambda}t + k}{2}\right)$$

випливає

$$f_{\sum_{i=1}^Y X_i/\sqrt{\lambda}}(t) \leq \sum_{k=1}^{\infty} P(Y = k) \frac{\sqrt{\lambda}}{2} \cdot \frac{C}{\sqrt{k}} = D \sum_{k=1}^{\infty} P(Y = k) \sqrt{\frac{\lambda}{k}}.$$

Звідси отримуємо

$$f_{\sum_{i=1}^Y X_i/\sqrt{\lambda}}(t) \leq D \mathbb{E}\left[\left(\frac{\lambda}{Y}\right)^{1/2}\right].$$

Згідно з лемою 3.4, маємо

$$\mathbb{E}\left[\left(\frac{\lambda}{Y}\right)^{1/2}\right] \rightarrow 1.$$

Отже

$$\sup_t f_{\sum_{i=1}^Y X_i/\sqrt{\lambda}}(t) \leq \tilde{D},$$

де  $\tilde{D} > 0$  — деяка константа.

Зауважимо, що

$$\begin{aligned} & \sup_x \left| P \left( \frac{S_Y(A)}{\sqrt{\lambda}} \leq x \right) - \Phi_{0,1}(x) \right| \leq \\ & \leq \sup_x \left| P \left( \frac{S_Y(A)}{\sqrt{\lambda}} \leq x \right) - P \left( \frac{S_Y}{\sqrt{\lambda}} \leq x \right) \right| + \sup_x \left| P \left( \frac{S_Y}{\sqrt{\lambda}} \leq x \right) - \Phi_{0,1}(x) \right|. \end{aligned} \quad (3.3)$$

Оцінка для другого доданку в правій частині нерівності (3.3) є відомою [18]. За аналогом нерівності Беррі-Ессеєна для випадкових сум похибка наближення суми незалежних випадкових величин нормальним розподілом оцінюється як

$$\left| \mathbb{P} \left( \frac{S_Y}{\sqrt{\lambda}} \leq x \right) - \Phi(x) \right| = O \left( \frac{1}{\sqrt{\lambda}} \right).$$

Введемо відстань Колмогорова:

$$d_K = \sup_x \left| \mathbb{P} \left( \frac{S_Y(A)}{\sqrt{\lambda}} \leq x \right) - \mathbb{P} \left( \frac{S_Y}{\sqrt{\lambda}} \leq x \right) \right|.$$

Для відстані Колмогорова  $d_K(U, V)$  між в.в.  $U, V$  справедливою є оцінка [16]

$$d_K(U, V) \leq \sqrt{2C \mathbb{E}|U - V|},$$

де  $C = \sup_{x \in \mathbb{R}} f_V(x)$ , а  $f_V$  — щільність розподілу випадкової величини  $V$ .

Нехай

$$U = \frac{S_Y(A)}{\sqrt{\lambda}}, \quad V = \frac{S_Y}{\sqrt{\lambda}}.$$

Використовуючи вищенаведену оцінку для щільності випадкової величини  $V$ , отримуємо:

$$d_K(U, V) \leq \sqrt{\tilde{D} \cdot \mathbb{E}|U - V|}.$$

Оскільки

$$U - V = \frac{1}{\sqrt{\lambda}} \sum_{i=1}^Y X_i \mathbf{1}(i \notin A),$$

то

$$\begin{aligned}
 d_K(U, V) &\leq \sqrt{\frac{\tilde{D}}{\sqrt{\lambda}} \mathbb{E} \left| \sum_{i=1}^Y X_i \mathbf{1}(i \notin A) \right|} \\
 &\leq \sqrt{\frac{\tilde{D}}{\sqrt{\lambda}} \mathbb{E} \left( \sum_{i=1}^Y \mathbf{1}(i \notin A) \right)} \\
 &\leq \sqrt{\frac{\tilde{D}}{\sqrt{\lambda}} \mathbb{E}[Y - d_Y(A)]} \\
 &\leq \sqrt{\frac{\tilde{D}}{\sqrt{\lambda}} \mathbb{E}(Y^\alpha)} \\
 &\leq \sqrt{\frac{\tilde{D}}{\sqrt{\lambda}}} \lambda^\alpha = C_1 \lambda^{\frac{1}{2}(\alpha - \frac{1}{2})},
 \end{aligned}$$

Отже, для оцінки відстані від нормального розподілу отримуємо

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \frac{S_Y(A)}{\sqrt{\lambda}} \leq x \right) - \Phi(x) \right| \leq \frac{C_1}{\sqrt{\lambda}^{(\frac{1}{2}-\alpha)}} + \frac{C_2}{\sqrt{\lambda}}$$

□

**Зауваження.** Відмітимо, що

$$\frac{C_1}{\sqrt{\lambda}^{(\frac{1}{2}-\alpha)}} + \frac{C_2}{\sqrt{\lambda}} \rightarrow \begin{cases} 0, & \text{якщо } \alpha < \frac{1}{2} \\ \infty, & \text{якщо } \alpha > \frac{1}{2}. \end{cases}$$

Отже, при  $\alpha > \frac{1}{2}$  отримана оцінка не дозволяє вловити швидкість збіжності прорідженої випадкової суми до нормального розподілу.

### 3.6 Оцінювання параметра прорідження $\beta$

У цьому розділі розглядається задача статистичного оцінювання параметра  $\beta$ , який визначає структуру множини прорідження.

Розглянемо

$$A_\beta = \mathbb{N} \setminus \{[k^\beta] : k \geq 1\}, \quad \beta > 1.$$

Тоді дисперсія прорідженої суми має вигляд

$$\text{Var}(S_Y(A_\beta)) = d_Y(A_\beta),$$

де  $d_Y(A_\beta)$  — кількість збережених індексів після прорідження.

Для даного класу множин виконується асимптотика

$$d_Y(A_\beta) \sim \lambda - \lambda^{1/\beta}, \quad \lambda \rightarrow \infty.$$

Звідси

$$\mathbb{E}[S_Y(A_\beta)^2] \sim \lambda - \lambda^{1/\beta}.$$

Нехай експеримент повторюється  $M$  разів і спостерігаються значення

$$S_1(A_\beta), S_2(A_\beta), \dots, S_M(A_\beta).$$

Введемо вибірковий другий момент

$$\overline{S(A_\beta)^2} = \frac{1}{M} \sum_{i=1}^M S_i(A_\beta)^2.$$

Метод моментів базується на порівнюванні теоретичного та вибіркового моментів:

$$\overline{S(A_\beta)^2} \approx \lambda - \lambda^{1/\beta}.$$

Звідси отримуємо оцінку параметра  $\beta$ :

$$\hat{\beta}_{\text{МоМ}} = \frac{\log \lambda}{\log \left( \lambda - \overline{S(A_\beta)^2} \right)}.$$

### 3.7 Результати моделювання

У цьому підрозділі наведено результати чисельних експериментів, метою яких є аналіз впливу параметра  $\beta$  на відповідні статистичні характеристики досліджуваних конструкцій.

#### Кількість виключених індексів.

Для дослідження структури множин прорідження для кожного значення

$$\lambda \in \{100, 10^4, 10^6\}$$

генерувалася випадкова кількість індексів

$$n \sim \text{Poisson}(\lambda).$$

Далі для параметрів

$$\beta \in \{1.2, 1.5, 1.8, 2, 3, 4, 5, 6\}$$

будувалися множини  $A_\beta$ , які формувалися шляхом виключення індексів вигляду

$$\lfloor k^\beta \rfloor, \quad k = 1, 2, \dots$$

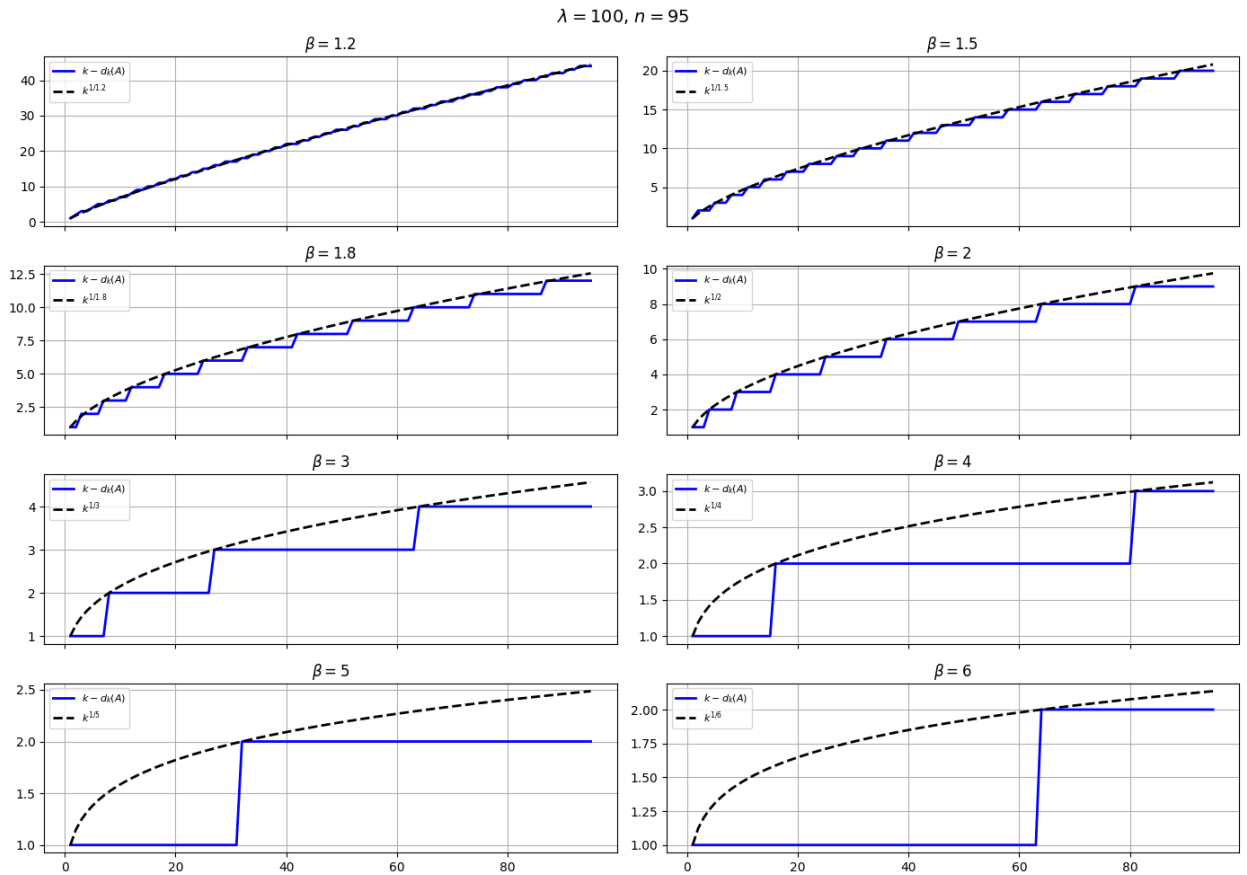
Для аналізу властивостей таких множин розглядалася функція

$$k - d_k(A),$$

де  $d_k(A)$  позначає кількість елементів множини  $A$ , що не перевищують  $k$ . Також для порівняння з теоретичними результатами будувалася функція

$$k^{1/\beta},$$

яка описує асимптотичний порядок зростання кількості виключених індексів.

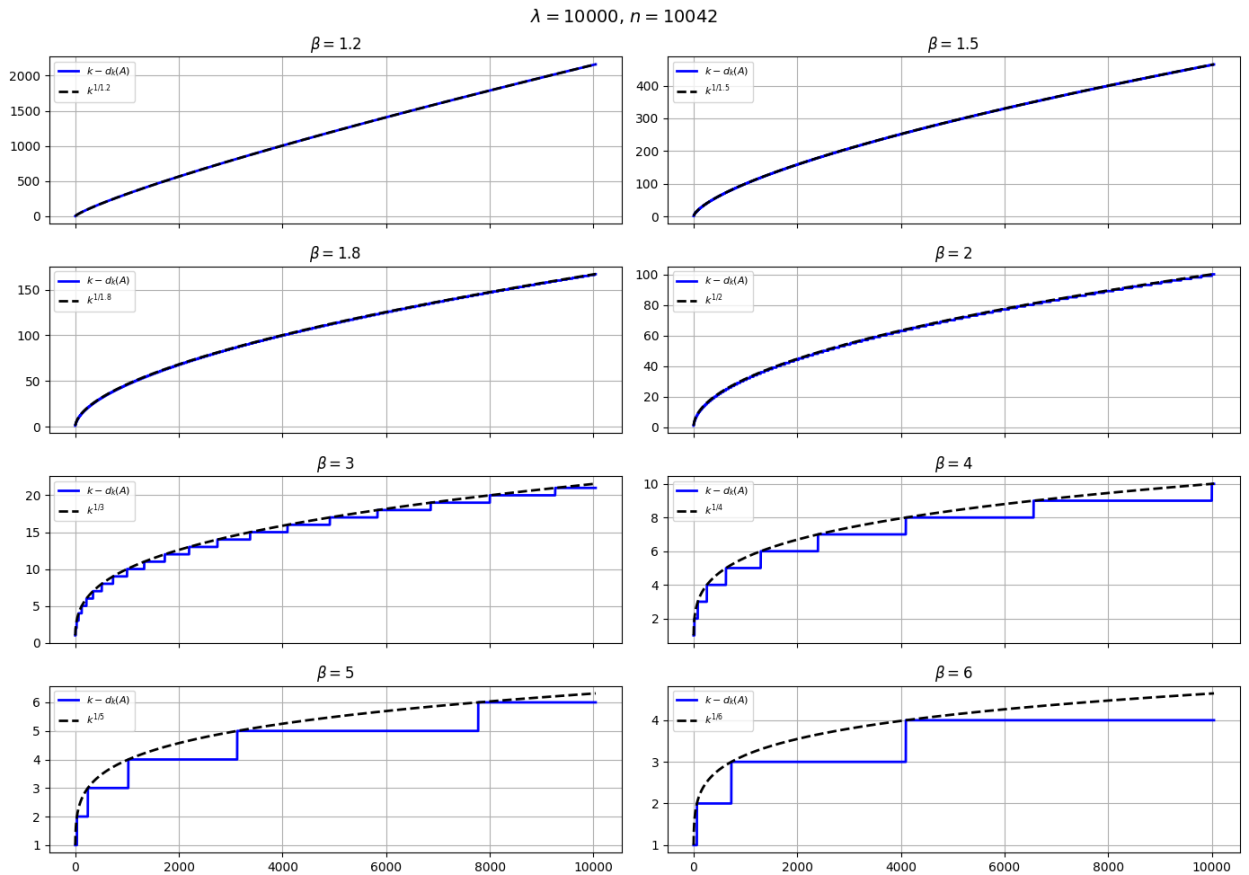


**Рисунок 3.1** – Кількість виключених індексів для  $\lambda = 100$ .

На рис. 3.1 представлено результати моделювання для початкового значення інтенсивності  $\lambda = 100$  ( $n = 95$ ). При такому відносно невеликому обсязі вибірки чітко проявляється дискретний характер досліджуваного процесу.

Для малих значень параметра прорідження ( $\beta = 1.2$  та  $\beta = 1.5$ ) емпірична функція  $k - d_k(A)$  демонструє високу інтенсивність вилучення елементів і візуально наближається до теоретичної кривої  $k^{1/\beta}$ .

Проте, починаючи з  $\beta \geq 1.8$  і особливо для  $\beta \in \{3, 4, 5, 6\}$ , графік набуває вираженої ступінчастої структури. Через швидке зростання відстані між сусідніми індексами вигляду  $[k^\beta]$  функція тривалий час залишається сталою, утворюючи широкі горизонтальні плато («сходи»), які дискретно піднімаються на одиницю лише в точках виключення. Кількість вилучених індексів на всьому проміжку при  $\beta = 6$  не перевищує двох елементів.

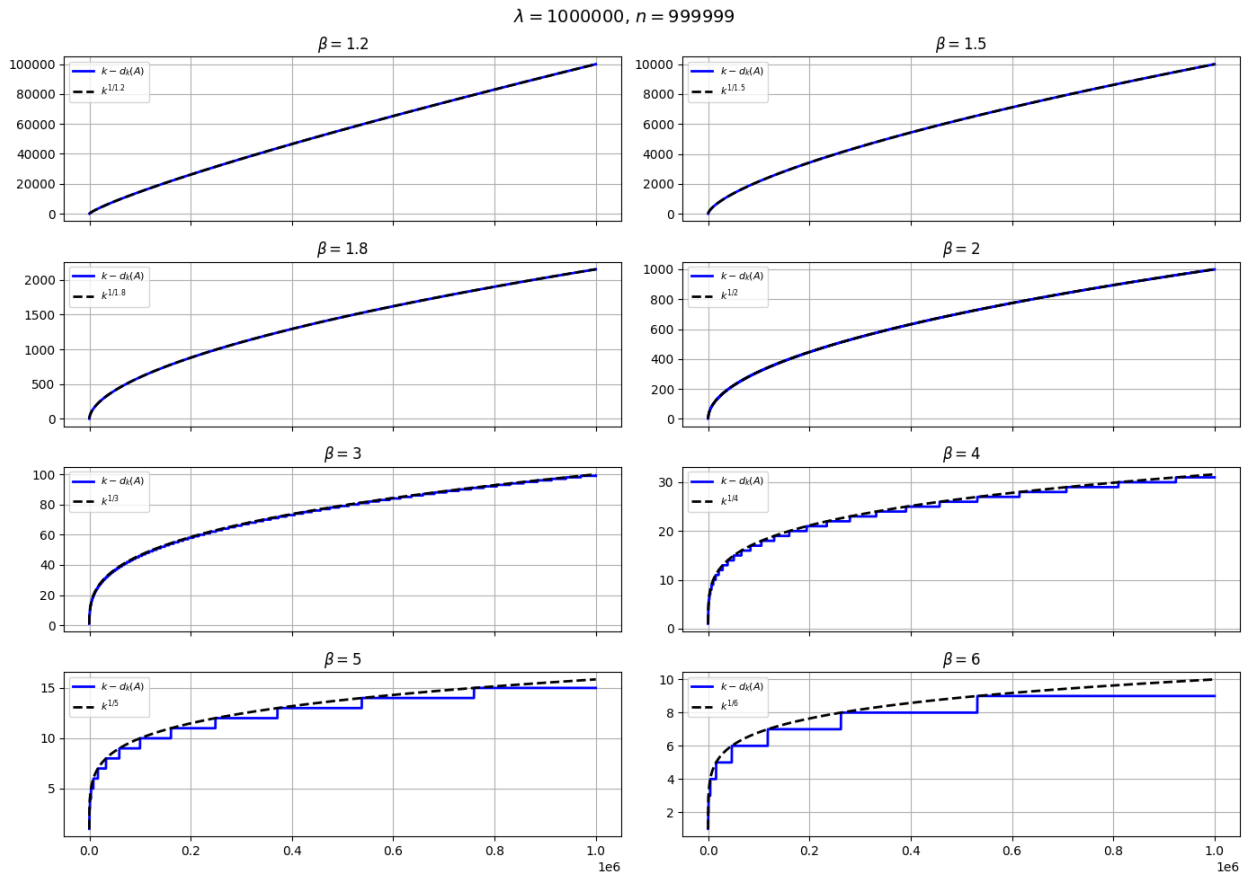


**Рисунок 3.2** – Кількість виключених індексів для  $\lambda = 10\,000$ .

Зі збільшенням інтенсивності до  $\lambda = 10\,000$  ( $n = 10\,042$ ), як показано на рис. 3.2, спостерігається суттєва трансформація вигляду кривих за рахунок масштабного розширення осі абсцис.

Для параметрів  $\beta \leq 2$  дискретні кроки емпіричної функції стають майже нерозрізними, а синя суцільна лінія практично ідеально зливається з теоретичною штриховою огинаючою  $k^{1/\beta}$ . Це експериментально засвідчує високу точність асимптотичного наближення на помірних інтервалах.

Водночас для великих значень параметра ( $\beta = 5$  та  $\beta = 6$ ) ступінчаста структура все ще залишається добре помітною на графіках, оскільки щільність виключених індексів для таких степенів лишається низькою, а максимальна кількість вилючених елементів наприкінці проміжку становить лише 4–6 індексів.



**Рисунок 3.3** – Кількість виключених індексів для  $\lambda = 1\,000\,000$ .

На рис. 3.3 відображено поведінку системи за умов великої вибірки з інтенсивністю  $\lambda = 1\,000\,000$  ( $n = 999\,999$ ).

У цьому масштабі для більшості досліджуваних параметрів ( $\beta$  від 1.2 до 3) випадкові відхилення повністю нівелюються, а графік емпіричної кількості вилучених індексів перетворюється на абсолютно гладку криву, що бездоганно повторює степеневий тренд  $k^{1/\beta}$ .

Ефект ступінчасті плато зберігаються лише на підграфіках для  $\beta = 5$  та  $\beta = 6$ . Навіть на інтервалі в мільйон точок кількість виключених елементів для  $\beta = 6$  досягає лише 9, що наочно ілюструє, наскільки сильно уповільнюється процес прорідження зі зростанням показника степеня.

Також на всіх графіках помітно, що теоретична крива завжди діє як точна верхня межа, що зумовлено математичною специфікою операції взяття нижньої цілої частини.

## Повна та проріджена випадкова суми.

Для дослідження впливу прорідження на поведінку випадкових сум для кожного значення

$$\lambda \in \{100, 10^4, 10^6\}$$

генерувалася випадкова величина

$$n \sim \text{Poisson}(\lambda),$$

а також послідовність незалежних однаково розподілених випадкових величин

$$X_i \sim U(-1, 1).$$

Для кожного значення параметра

$$\beta \in \{1.2, 1.5, 1.8, 2, 3, 4, 5, 6\}$$

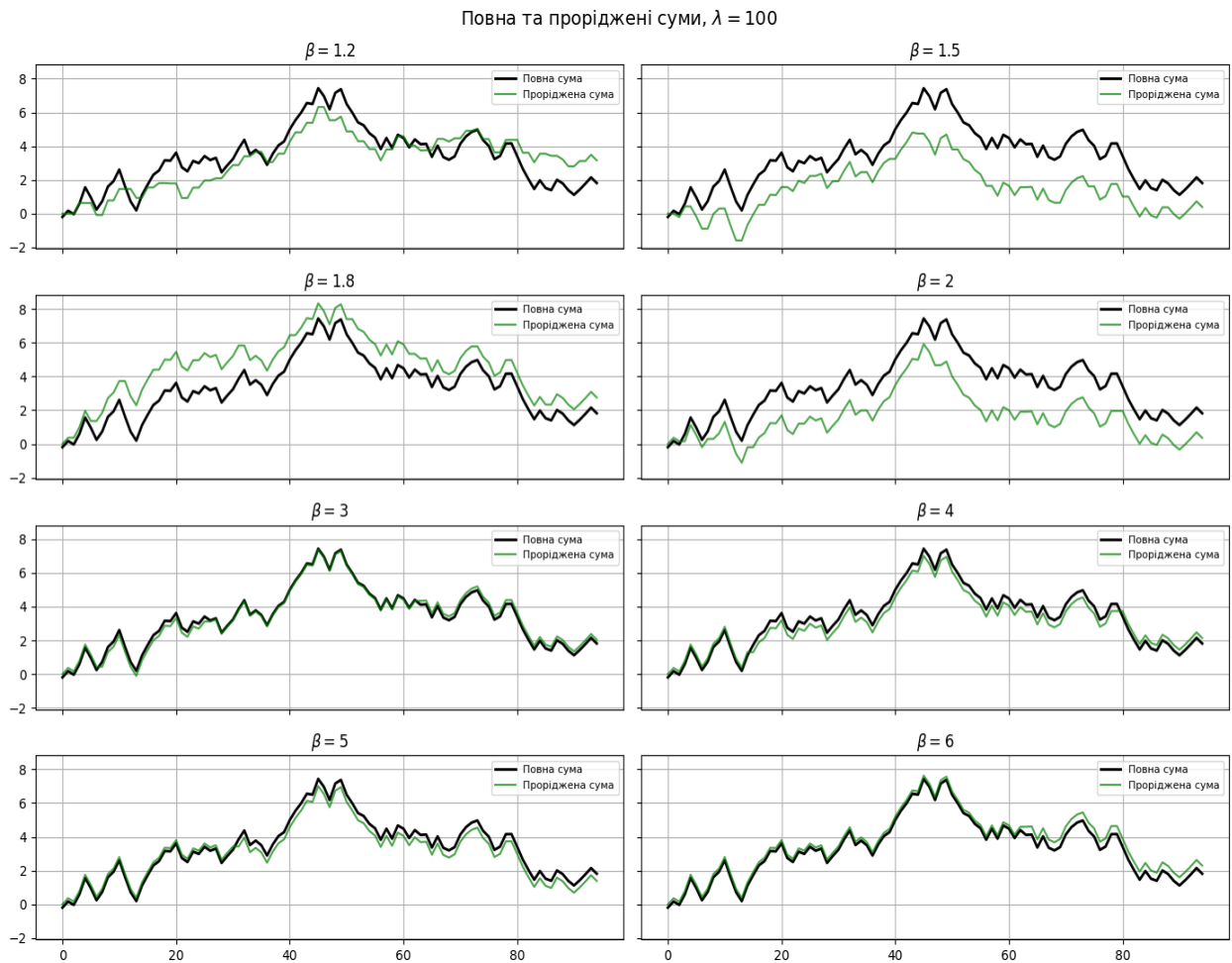
будувалася множина прорідження  $A_\beta$ , після чого обчислювалися повна сума

$$S_n = \sum_{j=1}^n X_j$$

та проріджена сума

$$S_n(A_\beta) = \sum_{j \in A_\beta} X_j.$$

Побудовані графіки дозволяють порівняти поведінку повних і проріджених сум при різних параметрах прорідження.



**Рисунок 3.4** – Порівняння повної та прорідженої сум для  $\lambda = 100$ .

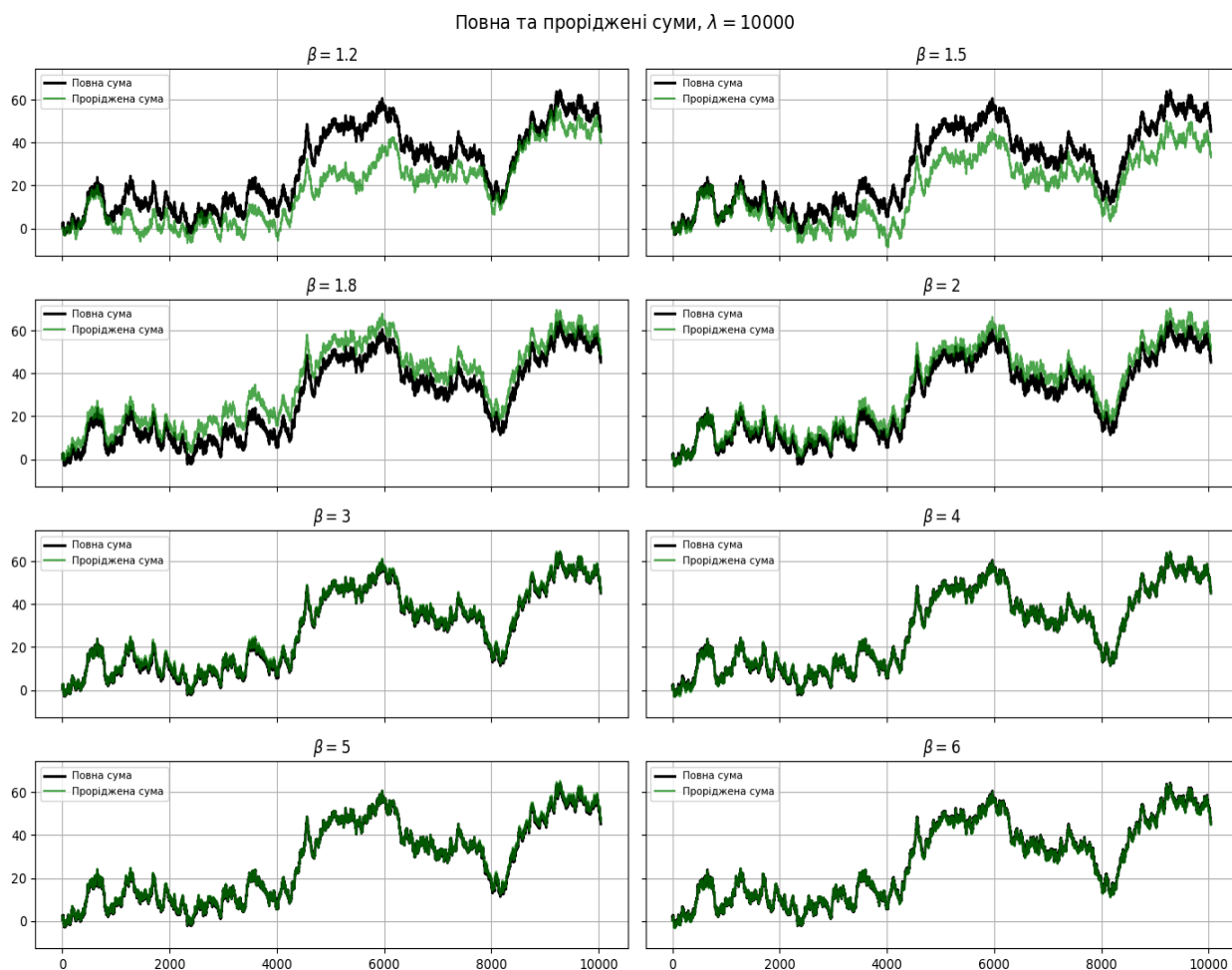
На рис. 3.4 представлено порівняльні траєкторії повної та прорідженої випадкових сум для малої інтенсивності  $\lambda = 100$ . Чорна крива відображає еволюцію повної суми  $S_n$ , тоді як зелена крива відповідає прорідженій сумі  $S_n(A_\beta)$ .

Через невелику довжину випадкового блукання вплив оператора вилучення індексів  $[k^\beta]$  є максимально дискретним та рельєфним.

При малих значеннях параметра ( $\beta = 1.2$  та  $\beta = 1.5$ ) з початкової послідовності вилучається значна частка доданків. Оскільки випадкові величини мають симетричний рівномірний розподіл  $U(-1,1)$ , вилучення великої кількості елементів призводить до суттєвої та нерегулярної розбіжності між двома траєкторіями, що проявляється у вигляді фазових зсувів та амплітудних відхилень.

Цікаво, що при  $\beta = 1.8$  спостерігається ситуація, коли проріджена сума на певному інтервалі перевищує повну суму, що пояснюється випадковим вилученням переважно від'ємних значень  $X_i$ .

Зі зростанням параметра прорідження ( $\beta \geq 3$ ) кількість вилучених елементів стає незначною, і зелена крива практично повністю повторює чорну траєкторію з мінімальними локальними відхиленнями.



**Рисунок 3.5** – Траєкторії сум для  $\lambda = 10\,000$ .

Зі збільшенням масштабу системи до  $\lambda = 10\,000$ , як демонструють підграфіки на рис. 3.5, характер розбіжності траєкторій набуває вираженої стохастичної закономірності.

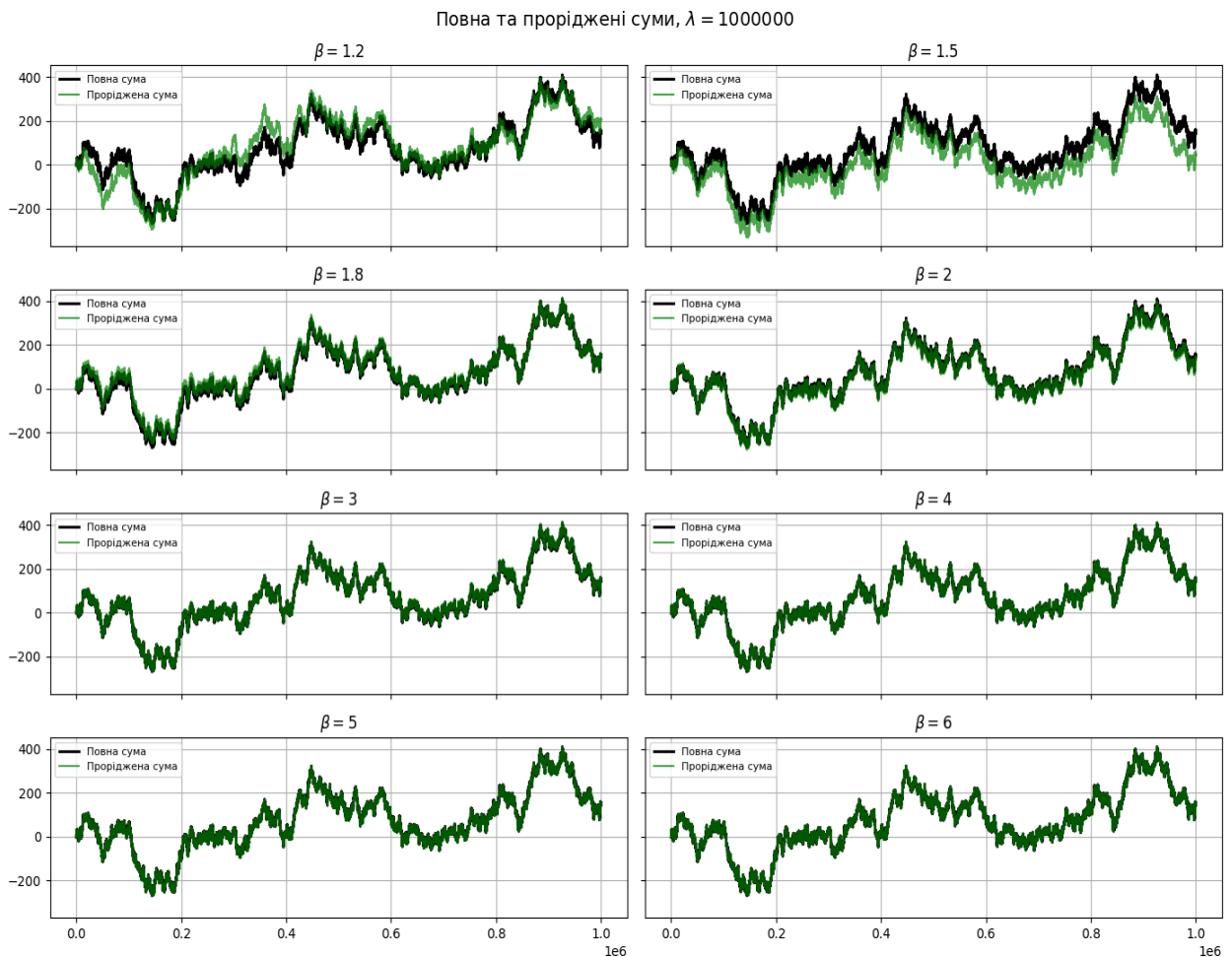
При малих степенях прорідження ( $\beta = 1.2$  та  $\beta = 1.5$ ) спостерігається системне розщеплення випадкових блукань. Траєкторія

прорідженої суми зберігає глобальну структуру повної суми, однак зазнає локального «стискання» амплітуди.

Це пояснюється зменшенням дисперсії внаслідок меншої кількості доданків.

Починаючи з  $\beta = 2$ , відхилення між процесами починає швидко зменшуватися. Для  $\beta \in \{3, 4, 5, 6\}$  візуальна різниця між траєкторіями практично зникає, і процеси стають майже ідентичними.

Таким чином, при фіксованому  $\lambda$  прорідження з великим  $\beta$  не змінює суттєво асимптотичну поведінку суми.



**Рисунок 3.6** – Траєкторії сум для  $\lambda = 1\,000\,000$ .

На рис. 3.6 відображено результати симуляції для великої інтенсивності  $\lambda = 1\,000\,000$ .

Головною особливістю є те, що траєкторії повної та прорідженої сум практично повністю накладаються для всіх значень  $\beta \geq 1.8$ .

Відхилення спостерігаються лише для  $\beta = 1.2$  та  $\beta = 1.5$ , і мають характер плавних коливань без різких стрибків.

Для  $\beta \in \{2, 3, 4, 5, 6\}$  графіки демонструють повний візуальний збіг.

Таким чином, експеримент підтверджує, що при великих  $\lambda$  центральна гранична теорема зберігається для прорідженої суми з тими ж нормуваннями, що і для повної суми.

### **Поведінка відстані Колмогорова.**

У цьому підрозділі досліджується швидкість збіжності розподілу нормованих проріджених сум до граничного нормального розподілу. Зокрема, для різних значень параметра  $\beta$  та інтенсивності  $\lambda$  обчислюється відстань Колмогорова

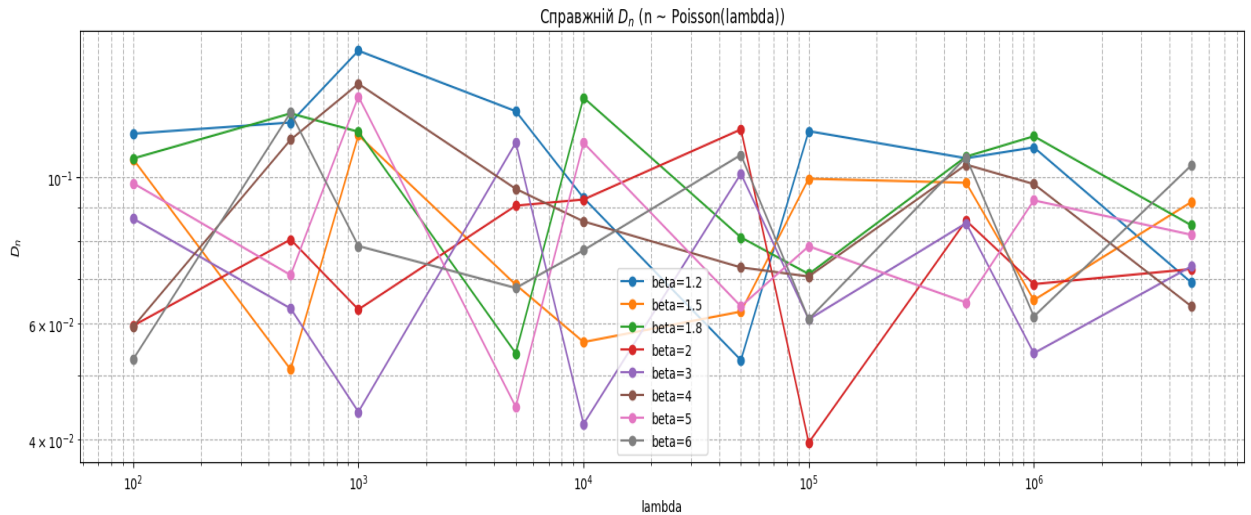
$$D_n = \sup_x |F_n(x) - \Phi(x)|,$$

де  $F_n(x)$  — емпірична функція розподілу нормованої прорідженої випадкової суми, а  $\Phi(x)$  — функція розподілу нормального розподілу з дисперсією  $1/3$ , що відповідає розподілу  $U(-1, 1)$ .

Для оцінки впливу параметра прорідження додатково розглядається масштабована величина

$$D_n \cdot \lambda^{1/2-1/\beta},$$

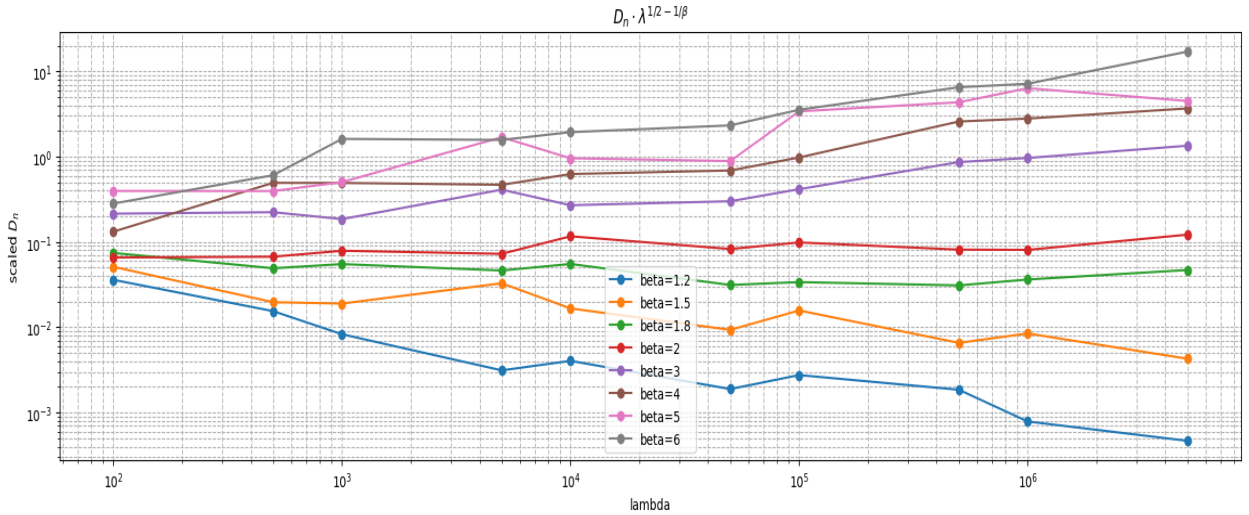
яка дозволяє перевірити характер передбачуваної асимптотичної залежності від  $\lambda$ . Графіки побудовано у логарифмічній шкалі для наочного порівняння поведінки статистики при різних значеннях параметрів.



**Рисунок 3.7** – Відстань Колмогорова  $D_n$  для проріджених сум  $S_Y(A)$  при різних  $\beta$ .

На рис. 3.7 представлено залежність початкової (немасштабованої) відстані Колмогорова  $D_n$  від інтенсивності  $\lambda$  для різних значень параметра прорідження  $\beta$ . Графіки побудовано у логарифмічному масштабі, що дозволяє виявити приховані степеневі закономірності та зменшити вплив випадкових коливань.

Спостерігається загальна тенденція до спадання  $D_n$  зі зростанням  $\lambda$ , що узгоджується з центральною граничною теоремою та збіжністю емпіричного розподілу до  $\Phi(x)$ . Водночас значення параметра  $\beta$  суттєво впливає на швидкість цього спадання. Для малих  $\beta$  (1.2 та 1.5) похибка зменшується швидко, тоді як для великих значень ( $\beta \geq 3$ ) спад є значно повільнішим. Це пояснюється тим, що при великому  $\beta$  кількість вилучених індексів є малою, і проріджена сума поводить себе майже як повна.



**Рисунок 3.8** – Нормована відстань  $D_n \cdot \lambda^{1/2-1/\beta}$ , що демонструє асимптотичну поведінку похибки.

Перехід до масштабованої відстані  $D_n \cdot \lambda^{1/2-1/\beta}$  (рис. 3.8) дозволяє перевірити теоретичний порядок збіжності. Отримані результати показують різну поведінку залежно від  $\beta$ .

Для критичного випадку  $\beta = 2$  масштабована крива є майже горизонтальною, що відповідає компенсації швидкості спадання. Для  $\beta < 2$  спостерігається спад масштабованих кривих, що вказує на швидшу фактичну збіжність. Натомість для  $\beta > 2$  криві зростають, причому чим більше  $\beta$ , тим сильніше зростання. Це пов'язано з тим, що множник  $\lambda^{1/2-1/\beta}$  у цьому випадку зростає і підсилює значення похибки.

Таким чином, результати узгоджуються з теоретичними передбаченнями та демонструють вплив параметра прорідження на швидкість збіжності.

### Висновки до розділу 3

У даному розділі досліджено асимптотичні властивості проріджених випадкових сум  $S_Y(A)$  при  $\lambda \rightarrow \infty$ . Отримано центральну граничну теорему для проріджених сум, яка встановлює їх збіжність до

нормального розподілу за відповідного нормування.

Показано асимптотичну еквівалентність повної та прорідженої сум у відповідних режимах параметрів, що дозволяє трактувати прорідження як стабільну операцію з точки зору граничної поведінки. Отримано оцінки швидкості збіжності до нормального розподілу.

Окрему увагу приділено задачі оцінювання параметра прорідження  $\beta$ . Запропоновано підхід, що базується на методі моментів спостережуваних сум.

Результати чисельного моделювання підтверджують теоретичні висновки: спостерігається узгодженість емпіричних розподілів із нормальним наближенням, а також коректна поведінка оцінок параметра  $\beta$  та відстані Колмогорова при зростанні  $\lambda$ .

## ВИСНОВКИ

У даній роботі досліджено асимптотичні властивості проріджених випадкових сум та вплив структури множини індексів на їх граничну поведінку.

У першому розділі було розглянуто основні підходи до вивчення випадкових сум та їх граничних властивостей, зокрема класичні результати центральної граничної теореми та оцінки Беррі–Ессеєна.

У другому розділі введено поняття натуральної щільності та проаналізовано її властивості для різних типів множин індексів. Досліджено взаємозв'язок між щільністю та поведінкою проріджених випадкових сум. Чисельні експерименти продемонстрували, що зі зменшенням щільності множини спостерігаються суттєві відхилення проріджених випадкових сум від повних, що підтверджує визначальну роль структури індексної множини.

У третьому розділі встановлено центральну граничну теорему для проріджених випадкових сум при  $\lambda \rightarrow \infty$ . Доведено асимптотичну еквівалентність повних та проріджених випадкових сум за відповідних умов, а також отримано оцінки швидкості збіжності до нормального розподілу. Окремо досліджено задачу оцінювання параметра прорідження  $\beta$  на основі моментного підходу. Результати чисельного моделювання підтвердили теоретичні висновки та показали узгодженість емпіричної поведінки з асимптотичними результатами.

Таким чином, у роботі показано, що натуральна щільність множини індексів є ключовим параметром, який визначає ступінь відхилення проріджених випадкових сум від класичної поведінки, а також суттєво впливає на їх граничні розподіли. Отримані результати узагальнюють класичні підходи теорії випадкових сум на випадок проріджених структур та можуть бути використані в задачах статистичного моделювання та аналізу даних із пропусками.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Grekos G. *Open Problems on Densities*. ResearchGate preprint, 2009.  
[https://www.researchgate.net/publication/318593875\\_Open\\_Problems\\_on\\_Densities](https://www.researchgate.net/publication/318593875_Open_Problems_on_Densities).
2. S. Gerhold, *Seminar Notes (plenary lecture)*, TU Wien. Available at:  
[https://fam.tuwien.ac.at/~sgerhold/pub\\_files/sem19/s\\_plenar.pdf](https://fam.tuwien.ac.at/~sgerhold/pub_files/sem19/s_plenar.pdf).
3. NYCU OpenCourseWare. *Berry–Esseen Theorem. Handout 13* <https://video.ocw.nycu.edu.tw/pub/apfc091/Handout/13.pdf>.
4. Shevtsova I. *On the absolute constants in the Berry–Esseen type inequalities for identically distributed summands*. <https://arxiv.org/abs/1111.6554>.
5. Delange H. *Introduction to Analytic and Probabilistic Number Theory* <https://www.ams.org/books/gsm/163/>.
6. Seppäläinen T. *Basics of Stochastic Analysis* <https://people.math.wisc.edu/~tseppalainen/courses/735/notes.pdf>.
7. Rotman J. J. *Advanced Modern Algebra* <https://www.anandinstitute.org/pdf/Advanced%20Modern%20Algebra%20-%20Joseph%20J.%20Rotman.pdf>.
8. Конспект лекцій. *Теорія ймовірностей* <https://dspace.lvduvs.edu.ua/bitstream/1234567890/629/1/%D1%82%D0%B5%D0%BE%D1%80%D1%96%D1%8F%20%D0%B9%D0%BC%D0%BE%D0%B2%D1%96%D1%80%D0%BD%D0%BE%D1%81%D1%82%D0%B5%D0%B9%20%D0%BF%D1%96%D0%B4%D1%80%D1%83%D1%87%D0%BD%D0%B8%D0%BA.pdf>.
9. Le Gall Y. *Probability Theory I. Lecture Notes, Fall 2021, University of Helsinki* <https://share.google/UJD48lQmp5qTLk5fs>.
10. Wang C., Wang J., Yan J. *Coherent States in Bernoulli Noise Functionals*. Bulletin of the Australian Mathematical Society, 2011.  
<https://www.cambridge.org/core/services/aop-cambridge-core/content/view/AEF7D3CAA38ADA3560474DA4443EAE53/S0004972711002139a.pdf/>

coherent\_states\_in\_bernoulli\_noise\_functionals.pdf.

11. Muroi Y. *Computation of Greeks Using the Discrete Malliavin Calculus and Binomial Tree*. Springer, 2022. <https://link.springer.com/book/10.1007/978-981-19-1073-9>.

12. Pitman J. *Lecture 14: Characteristic Functions and Continuity Theorem*, Stat 205, University of California, Berkeley. <https://www.stat.berkeley.edu/~pitman/s205f02/lecture14/lecture14.pdf>.

13. Stanford University — Stats 300B Lecture Notes, *Convergence of Random Variables*. Continuous mapping theorems, p. 13. <https://web.stanford.edu/class/stats300b/Slides/01-convergence.pdf>.

14. Shiryaev A.N. *Probability*. 1996. [https://ebooks.karbus.me/Mathematics/A.%20N.%20Shiryaev%20-%20Probability%20\(1996\)%20\[978-1-4757-2539-1\].pdf](https://ebooks.karbus.me/Mathematics/A.%20N.%20Shiryaev%20-%20Probability%20(1996)%20[978-1-4757-2539-1].pdf).

15. HandWiki. *Squeeze theorem*. [https://handwiki.org/wiki/Squeeze\\_theorem](https://handwiki.org/wiki/Squeeze_theorem).

16. Gaunt R., Li S. *Bounding Kolmogorov distances through Wasserstein and related integral probability metrics*. Journal of Mathematical Analysis and Applications, 2023. <https://www.sciencedirect.com/science/article/pii/S0022247X22009994>.

17. Jensen J. L. W. V. *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*. Acta Mathematica, 1906, Vol. 30, pp. 175–193. <https://projecteuclid.org/journalArticle/Download?urlId=10.1007%2FBBF02418571>.

18. Korolev V., Shevtsova I. *An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums*. arXiv preprint, 2009. <https://arxiv.org/pdf/0912.2795.pdf>.