**Міністерство освіти і науки України**

**Харківський національний університет ім. В. Н. Каразіна**

Факультет комп’ютерних наук

Кафедра

Моделювання систем і технологій

**Курсова робота**

на тему “Розподіл Гаусса — Кузьміна”

Виконав

студент групи КС-41

Моцний Данііл Вячеславович

Перевірив

Доктор фізико-математичних наук, професор кафедри моделювання систем і технологій

Лазурик Валентин Тимофійович

Харків – 2022

**ЗМІСТ**

[ВСТУП](#_heading=h.2y1tiscpvlp9) 3

[1 РОЗПОДІЛ](#_heading=h.xuce8pyocw3f) 4

[1.1 Розподіл Гаусса-Кузьміна](#_heading=h.mfg0r4s17lb0) 4

[1.2 Теорема Гаусса — Кузьміна](#_heading=h.lcf7yv4f8hn) 6

[1.3 Швидкість збіжності](#_heading=h.od9bqv3rhu4i) 6

[2 ВИЗНАЧЕННЯ](#_heading=h.6vxesvp9v63) 9

[3 МЕТОДИ ГЕНЕРУВАННЯ ВИПАДКОВОЇ ВЕЛИЧИНИ](#_heading=h.1rjbz5c0tfem) 9

[3.1 Метод зворотної функції](#_heading=h.87sqcf86artj) 9

[3.2 Алгоритм Метрополіса – Гастінгса](#_heading=h.mh9czvsfvmgk) 10

[4 ЗАСІБ ВИКОНАННЯ І РЕАЛІЗАЦІЇ](#_heading=h.w99y38vyr6w9) 11

[4.1 Структура проекту](#_heading=h.6u80nqh8jtqj) 11

[4.2 Зв’язок математичної моделі з програмною](#_heading=h.pss9o4ljgfb9) 12

[5 ОПИС ПРОГРАМИ, РОБОТА З ПРОГРАМОЮ](#_heading=h.rnj013q30ose) 13

[6 НАЛАШТУВАННЯ ТА ТЕСТУВАННЯ ПРОГРАМИ](#_heading=h.o9e3u3rwzo1y) 16

[ВИСНОВКИ](#_heading=h.1lv29owpyz5o) 18

[ВИКОРИСТАНІ ДЖЕРЕЛА](#_heading=h.s1llx4aom89p) 19

[ДОДАТОК А](#_heading=h.6nx8gvkyqu1j) 20

# ВСТУП

В цій роботі розглянуто розподіл Гаусса-Кузьміна — дискретний розподіл ймовірностей, який виникає як розподіл граничної ймовірності коефіцієнтів у розкладі випадкової величини.

Створено програмне забезпечення для моделювання розподілу, яке відповідає заданим критеріям:

1. Користувач повинен мати можливість зберігати результати моделювання наборів значень випадкової величини в файли в форматах доступних для подальшої обробки додатками MS Office.
2. Результати моделювання повинні бути представлені в графічному вигляді, який дозволяє порівняти закон розподілу ймовірності випадкової величини з результатами розрахунків, представленими у вигляді гістограм.
3. Інтегральні характеристики випадкової величини відповідні теоретичним значенням і значенням отриманим в результаті статистичного аналізу результатів розрахунку, повинні бути представлені в таблиці.
4. Окремими пунктами головного меню повинна бути допомога користувачеві, що містить теоретичні дані щодо розподілу та інструкцію по експлуатації програми, а також дані щодо розробника.

# 1 РОЗПОДІЛ

## 1.1 Розподіл Гаусса-Кузьміна

Розподіл Гаусса-Кузьміна — це розподіл появи натурального числа **k** у неперервній частці випадкового (або «загального») дійсного числа.

Розподіл названо на честь Карла Фрідріха Гаусса, який вивів його близько 1800 року, і Родіона Кузьміна, який дав обмеження на швидкість збіжності 1929 року.

У теорії ймовірностей і статистиці розподіл ймовірностей — це математична функція, яка визначає ймовірності появи різних можливих результатів експерименту. Це математичний опис випадкового явища в термінах його вибіркового простору та ймовірностей подій (підмножин вибіркового простору).

Натуральні числа — числа, що виникають природним чином при лічбі.

Дійсні числа - це будь-яке число, яке відповідає точці на дійсній прямій і може бути класифіковане на натуральне, ціле, раціональне та ірраціональне числа.

Розглянемо , визначене для дійсного числа x за допомогою

Так що є дробовою частиною від

Це можна визначити рекурсивно:

та

Де

та

просто n-й член ланцюгового дробу, x = [a0, a1, ...]

Розподіл дробового значення розглядався Гаусом у листі до Лапласа від 30 січня 1812 року. У цьому листі Гаус сказав, що він може довести простим аргументом, що якщо , це ймовірність того, що дробове значення < x для випадкового x, тоді

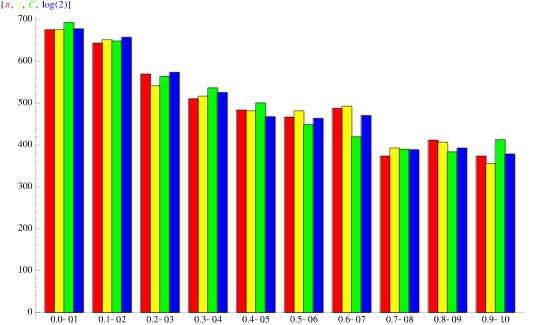


рис. 1 **—**  гістограма дробового значення ( ) для 5000 членів з різними константами

Вище наведено гістограму дробового значення ( ) для 5000 членів числа , гамма-константи Ейлера-Маскероні, константи K Каталона і натурального логарифма .

## 1.2 Теорема Гаусса — Кузьміна

Нехай

Нескінченний дріб розширення випадкового рівномірно розподіленого на (0, 1) числа *х*,

Аналогічно, нехай

тоді

прямує до нуля при n, що прямує до нескінченності.

## 1.3 Швидкість збіжності

Однак Гаусс не зміг описати поведінку поправочного члена, у

тому у 1928 році Кузьмін дав границю

з цього випливає що,

А також, у 1929 році Поль Леві її поліпшив ще раз з 0 < q < 1

з цього випливає що при q=0.7,

Едуард Вірсинг показав, що для *λ*=0.30366… (стала Гаусса — Кузьміна — Вірсинга), границя

існує для кожного , а функція *Ψ*(*х*) є аналітичною і задовольняє *Ψ*(0)=*Ψ*(1)=0.

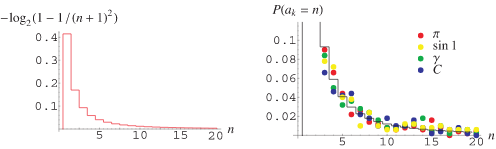


рис. 2 **—**  розподіл перших 500 доданків у неперервних частках з різними константами

Наведений вище графік показує розподіл перших 500 доданків у неперервних частках , sin(1), сталої Ейлера-Маскероні гамма та постійної C Коупленда-Ердеша.

Отже, розподіл був нормалізований належним чином, оскільки

# 2 ВИЗНАЧЕННЯ

Розподіл Гаусса-Кузьміна задається функцією ймовірності:

, де

( k довільне натуральне число)

Наприклад, для натурального k = 2,

# 3 МЕТОДИ ГЕНЕРУВАННЯ ВИПАДКОВОЇ ВЕЛИЧИНИ

## 3.1 Метод зворотної функції

Допустимо, що нам заданий інтегральний закон розподілу ймовірності F(x), де f(x) — функція щільності ймовірності й

Тоді досить розіграти випадкове число, рівномірно розподілене в інтервалі від 0 до 1. Оскільки функція F теж змінюється в даному інтервалі, то випадкова подія x можна визначити узяттям зворотної функції за графіком або аналітично: x = F –1 (r). Тут r — число, генероване еталонним ГВЧ в інтервалі від 0 до 1, x1 — згенерована в підсумку випадкова величина.

## 3.2 Алгоритм Метрополіса – Гастінгса

У статистиці та статистичній фізиці алгоритм Метрополіса – Гастінгса - це метод Монте-Карло Марковських ланцюгів (англ. MCMC) для отримання послідовності випадкових вибірок із таких розподілів, де звичайний прямий вибір є ускладненим.

# 4 ЗАСІБ ВИКОНАННЯ І РЕАЛІЗАЦІЇ

## 4.1 Структура проекту

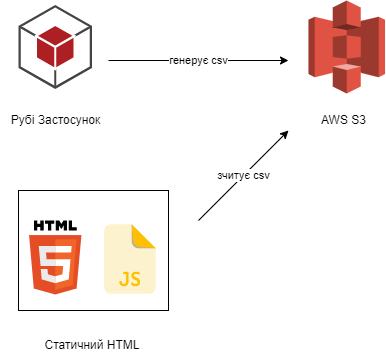
****

рис. 3 - структура взаємодії модулів ПЗ

На рисунку 3, зображено схему взаємодії застосунку. На першому етапі користувач виконує скрипт зі своїми параметрами, далі створюється csv файл, який він може переглянути. Файл csv автоматично завантажується на хмарне сховище Amazon S3. Після цього, користувач відкриває файл histogram.html та баче гістограму з данними файлу.

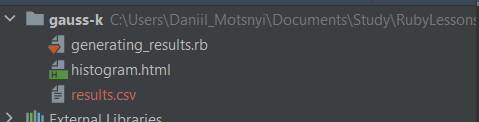
****

рис. 4 - структура проекту

Рисунок 4 відображає файли проекту, а також створений файл з результатами.

## 4.2 Зв’язок математичної моделі з програмною

Основою застосунку є математична функція закону розподілу ймовірності.

, де

Цій функції відповідає функція файлу generate\_results.rb

def gauss\_kuzmin\_distribution(*k*)

*# this func returns Gauss-Kuzmin distribution for selected k*

-Math.log((1 - (1 / ((1 + *k*) \* (1 + *k*)).to\_f)), 2)

end

За її допомогою, відбувається генерація csv файлу, який далі використовується гістограмою.

# 5 ОПИС ПРОГРАМИ, РОБОТА З ПРОГРАМОЮ

Розглянемо роботу застосунку на прикладі.

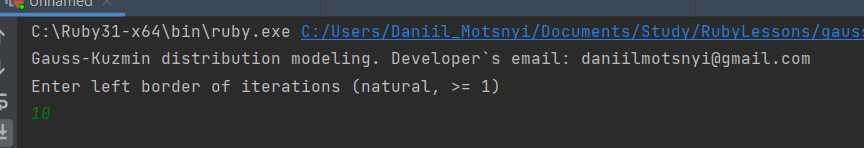
****

рис. 5 - введення користувачем даних

Після запуску скрипта generating\_results.rb, користувач має ввести ліву границю ітерацій для обчислення (рис. 5)

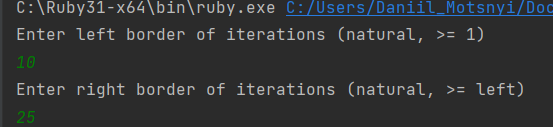
****

рис. 6 - введення користувачем даних

Далі, користувач має ввести праву границю ітерацій для обчислення (рис. 6)

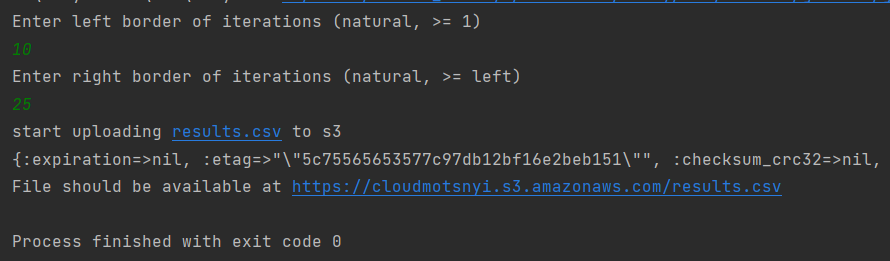
****

рис. 7 - результат виконання скрипту

Коли обчислення завершено, користувач отримує посилання на файл, може переглянути його вміст, а також гістограму.

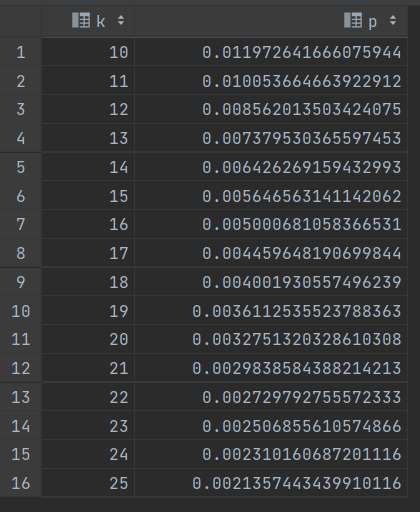


рис. 8 - csv файл з результатами

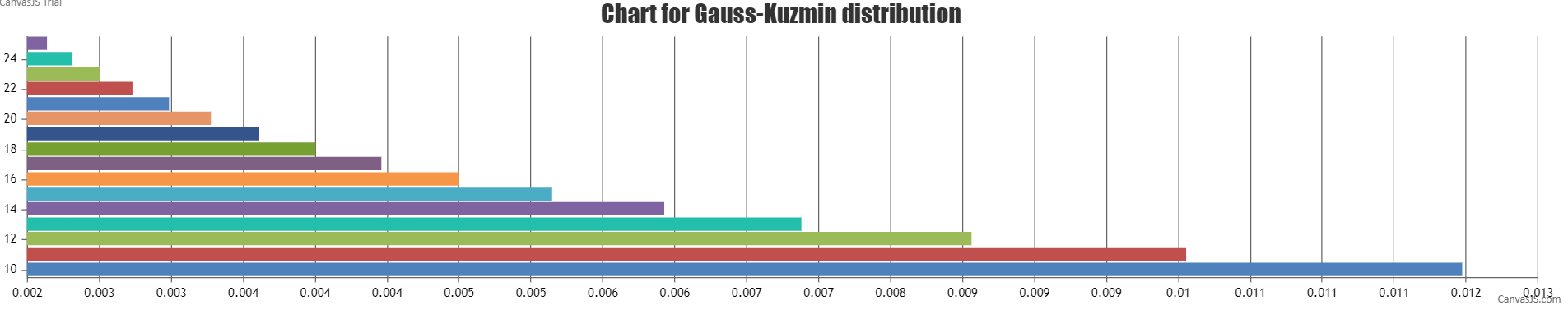


рис. 9 - гістограма

Для того щоб переглянути гістограму достатньо відкрити файл histogram.html

# 6 НАЛАШТУВАННЯ ТА ТЕСТУВАННЯ ПРОГРАМИ

Для налаштування програми, необхідно:

1. Інсталювати Ruby
2. Завантажити та відкрити проект
3. Виконати gem install aws-sdk-s3
4. Створити AWS аккаунт
5. Створити користувача у IAM та надати йому права
6. Створити S3 bucket
7. Надати йому такі параметри в поле CORS:

[

{

"AllowedHeaders": [

"\*"

],

"AllowedMethods": [

"GET",

"HEAD"

],

"AllowedOrigins": [

"\*"

],

"ExposeHeaders": [],

"MaxAgeSeconds": 3000

}

]

1. Вказати параметри користувача, назву та регіон S3 в змінні середовища
2. Запустити скрипт ruby generating\_results.rb

Для тестування застосунку, було обрано модель “Ad-Hoc”, це вид планового тестування програмного забезпечення, яке виконується без первинної документації тестового прикладу; однак спеціальне тестування також може бути застосоване до інших наукових досліджень і заходів з контролю якості.

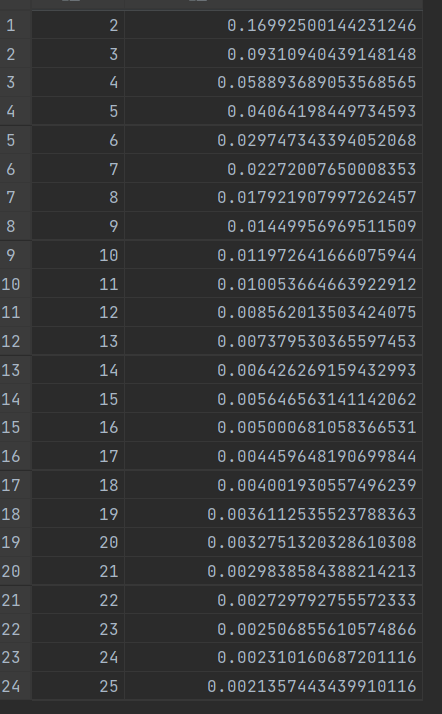
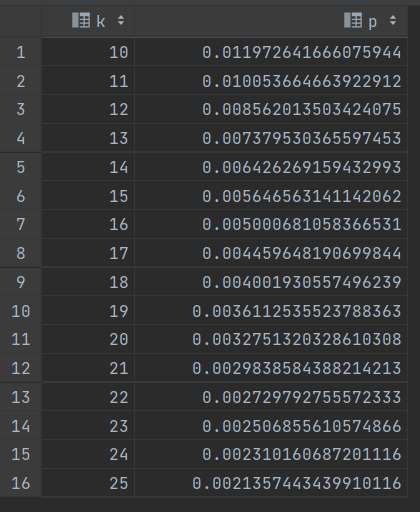


рис. 9 - тест проміжок 1-25

рис. 10 - тест проміжок 10-25

Далі наведено приклади файлів отриманих під час тестування, їх достовірність зумовлюється тим що при різних проміжках даних ми отримуємо однаковий результат для тих самих значень.

# ВИСНОВКИ

В результаті написання курсової роботи, на тему “Розподіл Гаусса — Кузьміна” було проведено дослідницьку роботу, детально описано розподіл ймовірностей появи натурального числа **k** у неперервній частці випадкового (або «загального») дійсного числа, пов’язану теорему Гаусса — Кузьміна, а також, швидкість збіжності. Створено програмне забезпечення, що дозволяє моделювати розподіл ймовірностей появи натурального числа, в залежності від заданих параметрів, переглянути файл з результатами обчислень функції вірогідностей, а також дослідити гістограму.

# ВИКОРИСТАНІ ДЖЕРЕЛА

* Kuzmin, R. O. (1928). "On a problem of Gauss". *Dokl. Akad. Nauk SSSR*: 375–380. (дата звернення: 23.11.2022).
* Nathan Ryan. Bucknell University. 2010. Statistics of the Gauss-Kuzmin Distribution. URL: <http://www.unix.bucknell.edu/~theburg/projects/reports/duff_poster.pdf> (дата звернення: 25.11.2022).
* Wikipedia. Gauss-Kuzmin Distribution. URL: <https://en.wikipedia.org/wiki/Gauss%E2%80%93Kuzmin_distribution> (дата звернення: 27.11.2022).
* Wikiland. Gauss-Kuzmin Distribution. URL: <https://www.wikiwand.com/en/Gauss%E2%80%93Kuzmin_distribution> (дата звернення: 01.12.2022).

# ДОДАТОК А

require 'aws-sdk-s3'

*ACCESS\_KEY\_ID* = ENV['ACCESS\_KEY\_ID']

*SECRET\_ACCESS\_KEY* = ENV['SECRET\_ACCESS\_KEY']

*REGION\_ID* = ENV['REGION\_ID']

*BUCKET\_NAME* = ENV['BUCKET\_NAME']

def gauss\_kuzmin\_distribution(*k*)

*# this func returns Gauss-Kuzmin distribution for selected k*

-Math.log((1 - (1 / ((1 + *k*) \* (1 + *k*)).to\_f)), 2)

end

begin

*file* = File.open('results.csv', 'w')

*# file.write("k,value\n")*

(0..20).each { |*\_*|

*k* = rand(1...1000)

*file*.write("#{*k*},#{gauss\_kuzmin\_distribution *k*}\n")

}

rescue IOError => *e*

puts *e*

ensure

*file*.close unless *file*.nil?

end

def upload\_2\_s3

*credentials* = Aws::Credentials.new(

ACCESS\_KEY\_ID,

SECRET\_ACCESS\_KEY

)

*s3* = Aws::S3::Client.new(

region: REGION\_ID,

credentials: *credentials*

)

File.open('results.csv', 'rb') do |*file*|

puts "start uploading results.csv to s3"

*resp* = *s3*.put\_object(bucket: BUCKET\_NAME, acl: "public-read", key: 'results.csv', body: *file*)

puts *resp*

end

puts "File should be available at https://#{BUCKET\_NAME}.s3.amazonaws.com/results.csv"

end