Національний технічний університет України
«Київський політехнічний інститут»
Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Прогнозування результатів експертного оцінювання рівня знань іноземної мови за допомогою сум випадкових величин

Виконав: **Гур'янов Олександр, гр. КА-21**

Науковий керівник: к. ф.-м. н., доц. **Каніовська І.Ю.**

2016 р.

- **Об’єкт дослідження**: процес оцінювання рівня знань студента експертною комісією.

- **Мета роботи**: розробка програмного продукту для знаходження ймовірностей отримання студентом кожної з можливих оцінок за виконання певного завдання на основі його особистих параметрів.

- **Метод дослідження**: висунута математична модель на основі розподілу суми випадкових величин.

- Побудована математична модель цілком базується на теорії ймовірностей та математичній статистиці, прогнозування за нею є альтернативою до використання методів регресійного аналізу або машинного навчання.

- Вона враховує те, що оцінка студента залежить від конкретної ситуації перевірки знань та процес допущення помилок є випадковим.

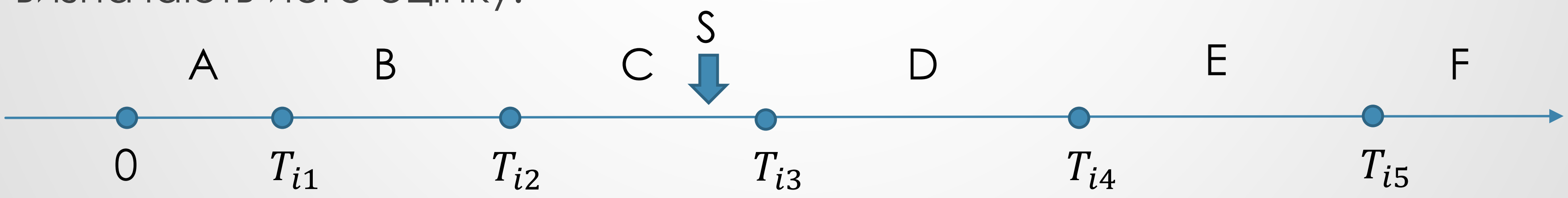
- Дану модель та створений на її основі програмний продукт можна використовувати в навчальному процесі для його моніторингу, підвищення ефективності та планування.

Актуальність дослідження

- У наш час активно розвиваються нові технології навчання та оцінювання знань.
- Великого поширення набуває дистанційна освіта, викладачі більше не мають можливості слідкувати за прогресом студентів. Необхідні нові способи оцінювання знань.
- У деяких завданнях з іноземних мов, наприклад, творах або усному мовленні, оцінювання традиційно проводиться експертами. Проте розвиток технологій обробки природної мови та розпізнавання мовлення дозволяє автоматизувати і ці процеси. Інструменти оцінювання враховують кількість і грубість допущених студентом помилок.
- Прогнозування оцінок студента на екзамені є важливою задачею.
- З усе більшим поширенням зовнішніх екзаменів, наприклад, IELTS або GMAT студенту необхідно оцінювати свою готовність до складання іспиту.
- З урахуванням прогнозних оцінок викладачі можуть коректувати навчальний процес.
- Тому необхідно розробляти методи прогнозування оцінок!

Постановка задачі

- Для оцінювання рівня знань студент виконує завдання, в якому може зробити певну кількість помилок.
- Якість виконання завдання студентом визначається певною випадковою величиною (ВВ) S, її розподіл специфічний для кожного студента.
- Виконання завдання оцінюють m екпертів $E_i,\ i = \overline{1,m}$.
- Студент отримує в результаті одну з k можливих оцінок.
- Кожному i-ому експерту відповідає система порогових значень $T_{ij},\ \mathrm{j} = \overline{1, \mathrm{k}-1}$, які визначають його оцінку.



- Студент отримує оцінку, яка враховує думки всіх експертів.
- **Необхідно** знайти імовірності, з якими студент отримає кожну з k оцінок.

Задача формування консолідованої експертної оцінки

- Розглядаються дві моделі оцінювання знань:
  - Залік (можливі 2 оцінки)
  - Екзамен (можливі k оцінок)
- Рішення експертів приймаються більшістю (модель «залік»), або консолідованою оцінкою вважається середнє арифметичне оцінок експертів (модель «екзамен»)
- Думки експертів можуть враховуватись з ваговими коефіцієнтами $w_i,\ i = \overline{1,m}$, якщо експерти нерівнозначні.

Для кожного випадку була виведена формула для імовірності отримання кожної можливої оцінки, що пов'язує ФР ВВ S та порогові значення експертів.

$$P(G_{j^*} = G_j) = \begin{cases} F_S\left(T_{\left\lfloor 1+\frac{m}{2}\right\rfloor}\right), & j = 1 \\ F_S\left(T_{\left\lfloor 1+mj-\frac{m}{2}\right\rfloor}\right) - F_S\left(T_{\left\lfloor 1+mj-\frac{3m}{2}\right\rfloor}\right), & 1 < j < k \\ 1 - F_S\left(T_{\left\lfloor 1+mk-\frac{3m}{2}\right\rfloor}\right), & j = k \end{cases}$$

Формула для моделі «екзамен»

Формалізація ВВ S

- Вважається, що є n типів помилок, які студент може допускати з певною частотою.
- Тривалість завдання вважається заданою.
- Кількість допущених помилок i-го типу X_i можна моделювати ВВ, що має розподіл Пуассона з параметром $\lambda_i t$, тобто $X_i \in Poiss(\lambda_i t)$, де
  - λ_i - середня кількість помилок i-го типу на одиницю часу
  - t – тривалість завдання

$$P(X_i = x) = \frac{(\lambda_i t)^x}{x!} e^{-\lambda_i t} \quad , \quad x = 0, 1, 2, \dots$$

- Кожному типу помилок присвоюється ваговий коефіцієнт θ_i, що визначає грубість.
- Якість виконання завдання визначається ВВ S – зваженою сумою пуассонівських ВВ:

$$S = \sum_{i=1}^{n} \theta_i X_i$$

Знаходження функції розподілу ВВ S

- Для того, щоб знайти імовірності отримання студентом кожної можливої оцінки, необхідно знати функцію розподілу ВВ S F_S (x):

$$F_S\,(\mathrm{x}) = \mathrm{P}(\mathrm{S} \leq x)$$

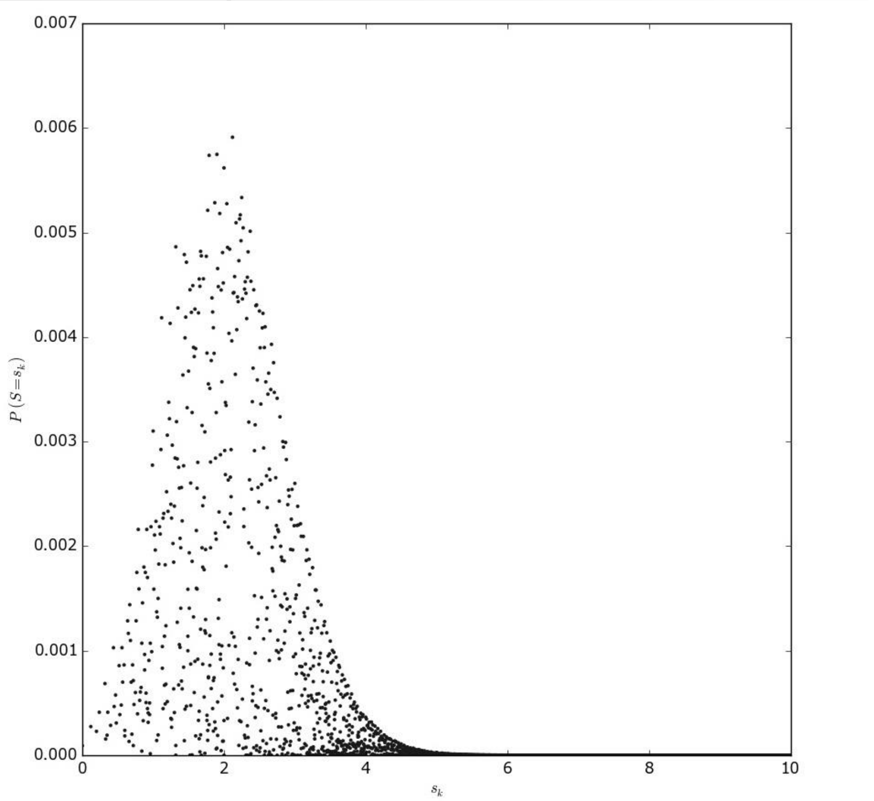
- Зважена сума пуассонівських ВВ є дискретною ВВ, але її розподіл є дуже нерегулярним, близьким до неперервного. Зручне аналітичне представлення імовірностей через параметри неможливе.

- Для знаходження точних значень ФР зваженої суми незалежних у сукупності пуассонівських ВВ запропоновано три методи:
  - Рекурентний
  - З використанням генератриси та символьної математики
  - З використанням характеристичної функції

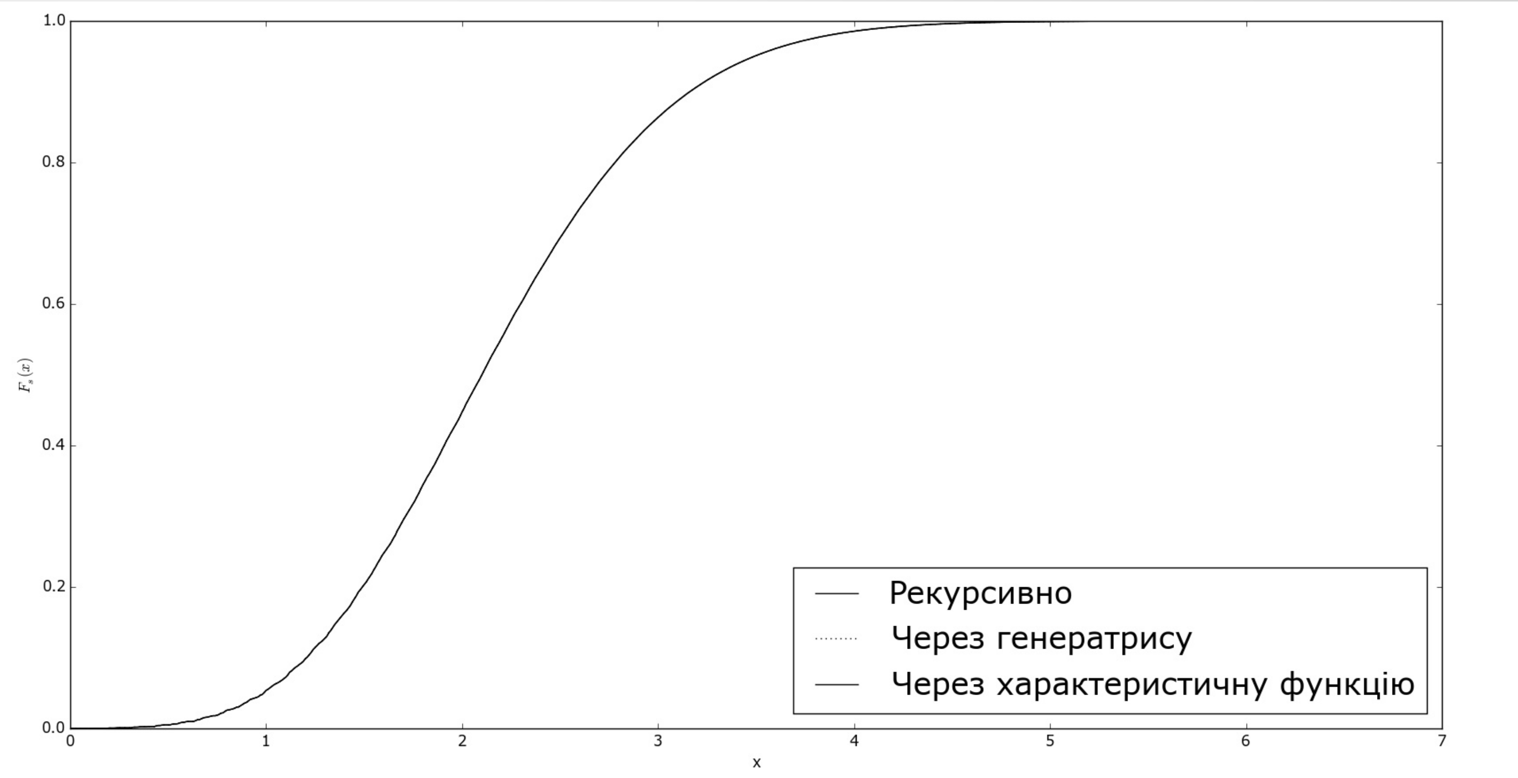
Приклад

$S = 0.35X_1 + 0.2X_2 + 0.12X_3 + 0.33X_4$

$X_1 \in Poiss(2.1)$, $X_2 \in Poiss(2.5)$,
$X_3 \in Poiss(3.0)$, $X_4 \in Poiss(1.7)$



Графік розподілу ВВ S

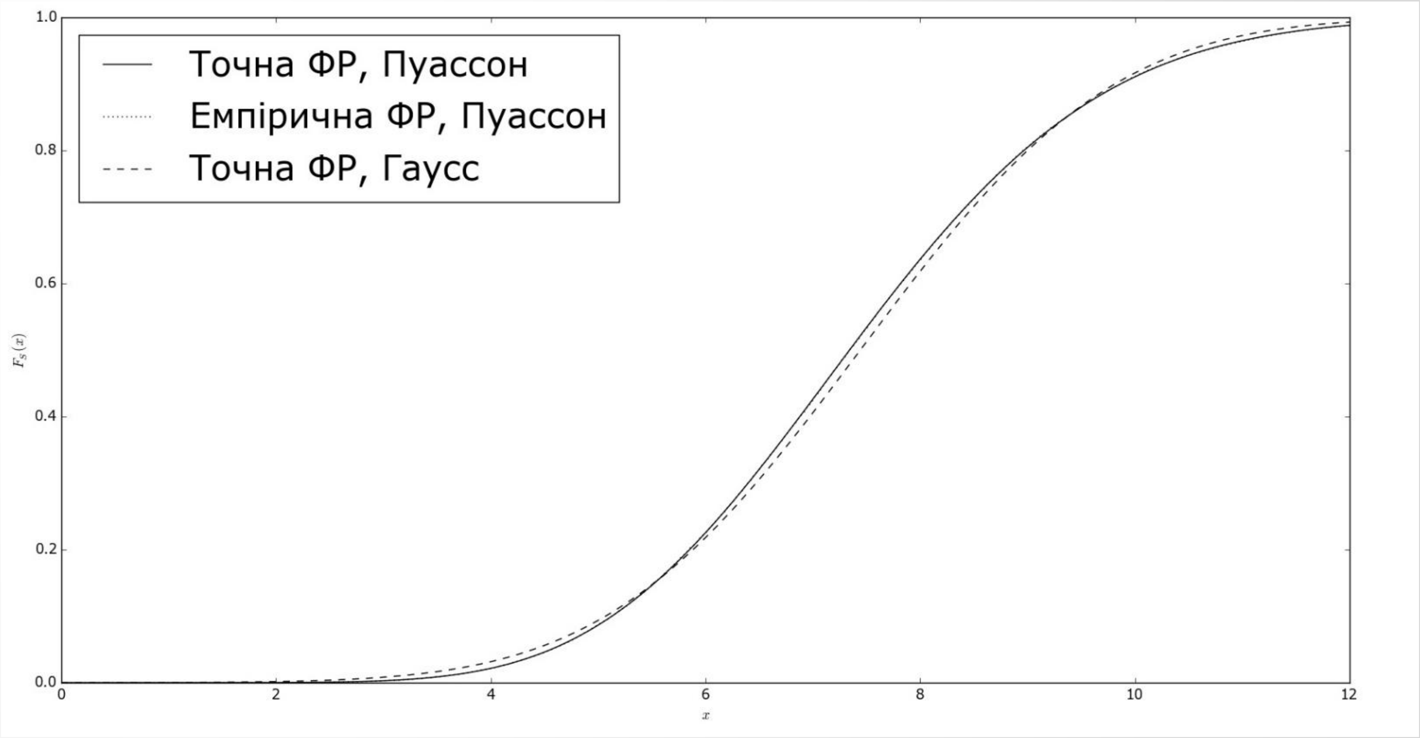


Функція розподілу ВВ S

- Час роботи методів: 10 с, 70 с, 15 с. При великому n оптимальним є метод з використанням характеристичних функцій.
- З графіка розподілу видно, що традиційна апроксимація нормальним розподілом буде неточною, оскільки чимало точок знаходиться під обвідною.

Апроксимація функції розподілу ВВ S

- При великій кількості доданків n методи знаходження точної функції розподілу працюють неприйнятно довго (750 с. при $n = 8$).
- Розглядається два варіанти апроксимації функції розподілу:
  - апроксимація емпіричною функцією розподілу (метод Монте-Карло)
  - апроксимація за допомогою нормального розподілу
- Апроксимація емпіричною функцією розподілу (обсяг вибірки 10000) є на порядок більш точною.



Точна функція розподілу ВВ S та варіанти апроксимації, $n = 8$

Оцінка параметрів

- Найкращою (незміщеною, консистентною, ефективною) точковою оцінкою параметра λ за вибіркою з K завдань з кількостями помилок $Y_1, Y_2, \dots, Y_K$ та тривалостями $t_1, \dots, t_K$ є:

$$\hat{\lambda} = \frac{\sum_{k=1}^{K} Y_k}{\sum_{k=1}^{K} t_k}$$

- Також була розглянута задача побудови довірчих інтервалів для математичного сподівання ВВ S за результатом єдиного спостереження без оцінювання параметрів λ_i. Проведено порівняння чотирьох методів побудови інтервалів:

1) Нормальні інтервали
2) ABC інтервали
3) ДКЕШ інтервали
4) Гамма-інтервали

1	2
3	4

$1-\alpha$ \ n	6		8		12	
0.95	92.72	95.32	92.8	95.29	93.9	95.21
	95.71	97.34	95.51	96.93	95.77	96.89
0.9	88.21	90.64	88.37	90.25	88.7	90.06
	91.71	93.96	91.57	93.61	91.18	92.91
0.8	79.2	80.32	79.35	0.8063	79.27	0.8
	83.67	86.97	83.37	86.7	82.47	84.73

Відсоток покриття мат. сподівання довірчим інтервалом (10000 симуляцій); n – кількість доданків у зваженій сумі

- Визначено, що найкращими є гамма-інтервали.

Випадок залежних пуассонівських ВВ

- Виходячи з конкретної задачі, ВВ X_i можуть бути залежними.
- Тоді випадковий вектор $\vec{X} = (X_1, X_2, \dots, X_K)^T$ має багатовимірний розподіл Пуассона.
- З функцією розподілу пуассонівського випадкового вектора незручно працювати, через це він мало застосовується.
- ВВ S є лінійним перетворенням випадкового вектора $\vec{X}$.
- Пропонується апроксимувати функцію розподілу ВВ S емпіричною функцією розподілу.
- Для цього генеруються реалізації пуассонівського випадкового вектора $\vec{X}$ з заданим центром розсіювання та коваріаційною матрицею за допомогою алгоритму Сіма.

Тестування алгоритму Сіма

- Було виконано генерацію 10000 реалізацій пуассонівського випадкового вектора $\vec{X}$ з параметрами $\vec{a} = (1{,}9;\ 2{,}5;\ 3{,}4;\ 4{,}32)$ та коваріаційною матрицею:

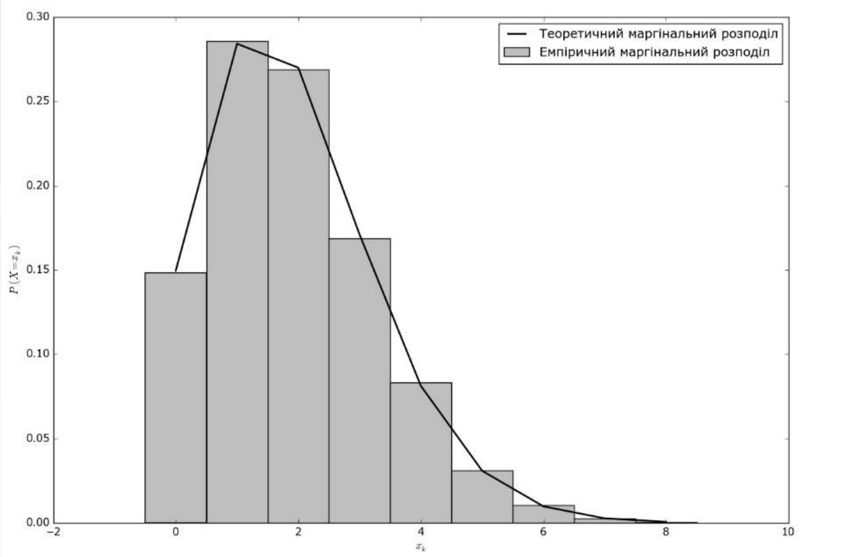
$$\Sigma = \begin{pmatrix} 1{,}9 & 0{,}9 & 0{,}8 & 1{,}2 \\ 0{,}9 & 2{,}5 & 0{,}4 & 0{,}6 \\ 0{,}8 & 0{,}4 & 3{,}4 & 0{,}8 \\ 1{,}2 & 0{,}6 & 0{,}8 & 4{,}32 \end{pmatrix}$$

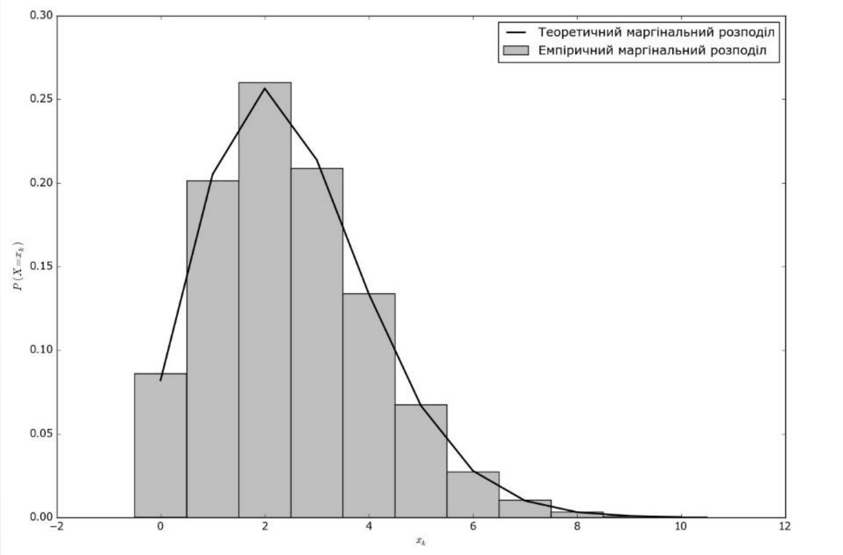
- Обчислені за отриманою вибіркою характеристики мають значення:

$$\hat{\vec{a}} = (1{,}9054;\ 2{,}494;\ 3{,}3933;\ 4{,}3071)$$

$$\hat{\Sigma} = \begin{pmatrix} 1{,}9084 & 0{,}9132 & 0{,}7837 & 1{,}1988 \\ 0{,}9132 & 2{,}5200 & 0{,}4637 & 0{,}6454 \\ 0{,}7837 & 0{,}4637 & 3{,}3902 & 0{,}8117 \\ 1{,}1988 & 0{,}6454 & 0{,}8117 & 4{,}2192 \end{pmatrix}$$

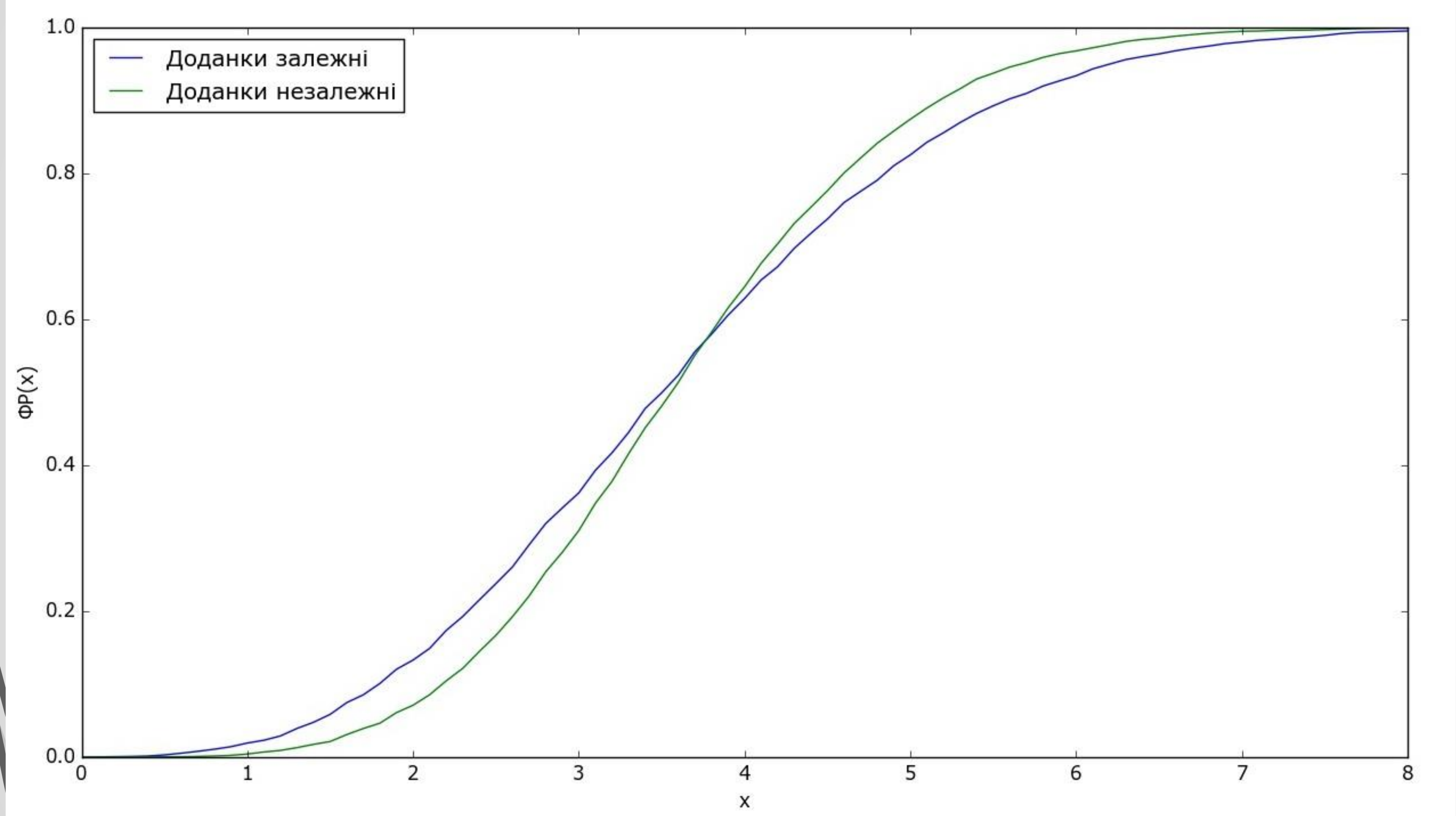
- Отримані емпіричні характеристики близькі до заданих!



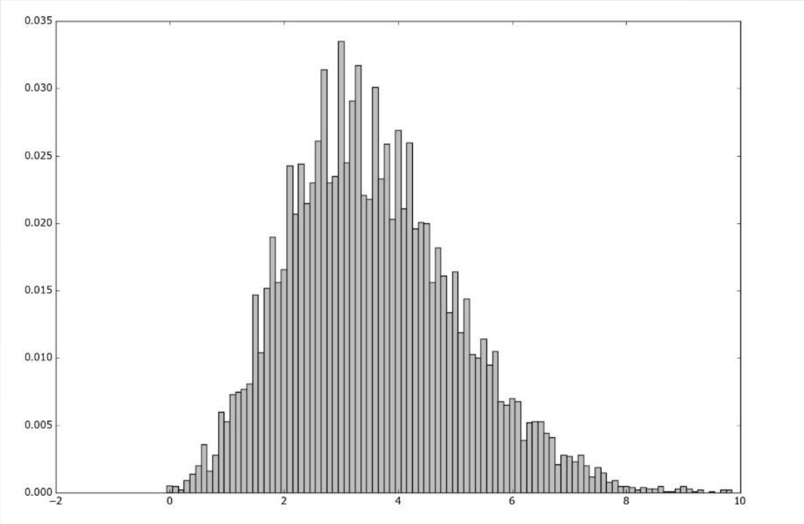


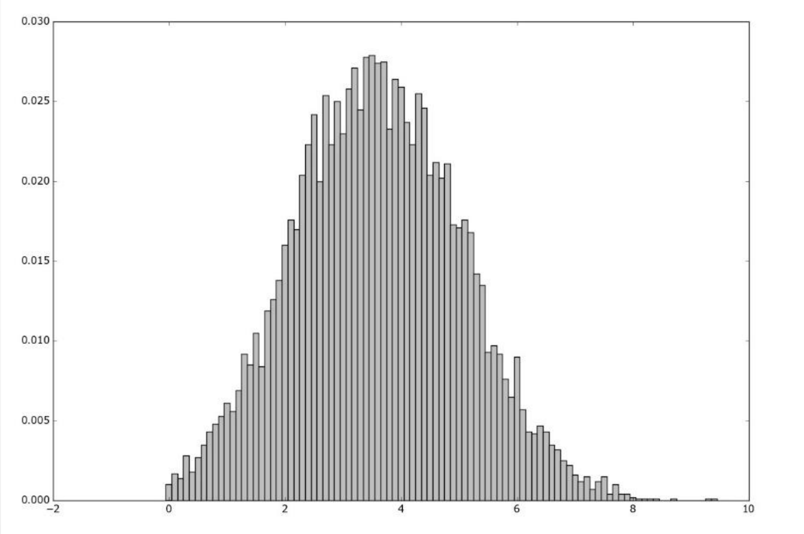
Графіки теоретичних та емпіричних маргінальних розподілів для 1-ої, 2-ої координат

Порівняння результатів



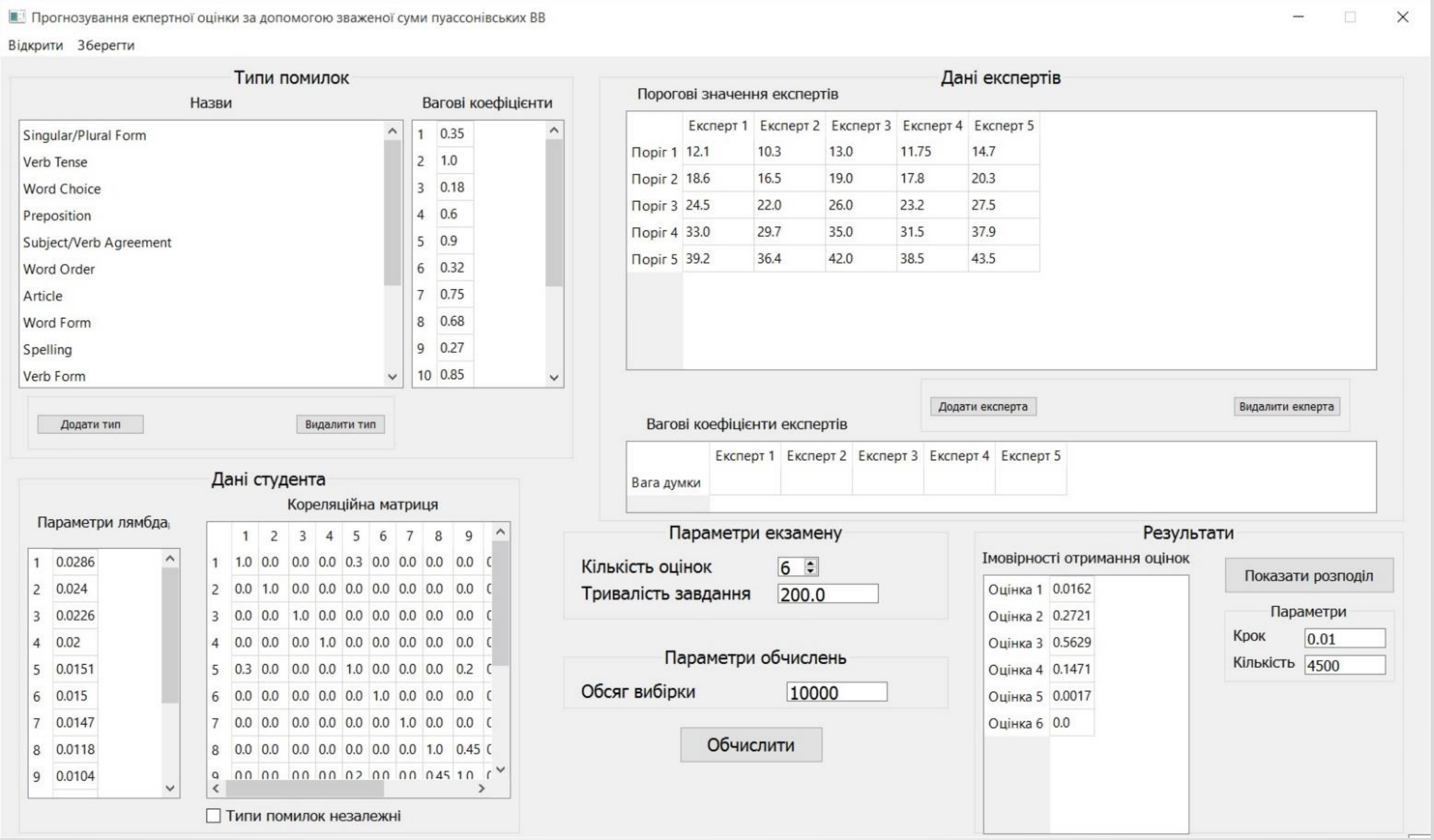
Функція розподілу ВВ S у випадках корельованих та некорельованих доданків з однаковим центром розсіювання вектора $\vec{X}$





Графіки емпіричного розподілу ВВ S на основі пуассонівського та гауссівського випадкового вектора

Програмний продукт



Головне вікно програми

Практичний приклад

- Побудовану модель було застосовано на прикладі з даними про написання творів з англійської мови малайзійськими студентами.

№	Назва типу помилки	Середня частота допущення, λ_i	Ваговий коефіцієнт, θ_i
1	Singular/Plural Form	0.0286	0.35
2	Verb Tense	0.024	1.0
3	Word Choice	0.0226	0.18
4	Preposition	0.02	0.6
5	Subject/Verb Agreement	0.0151	0.9
6	Word Order	0.015	0.32
7	Article	0.0147	0.75
8	Word Form	0.0118	0.68
9	Spelling	0.0104	0.27
10	Verb Form	0.0101	0.85
11	Capitalization	0.009	0.23
12	Wrong/Misused Word	0.0085	0.55
13	Missing word	0.0071	0.2
14	Redundancy	0.0054	0.1

Параметри типів помилок

- Твір оцінює комісія з 5 експертів.
- Обсяг твору: 200 слів.

	E_1	E_2	E_3	E_4	E_5
T_1	12,1	10,3	13,0	11,75	14,7
T_2	18,6	16,5	19,0	17,8	20,3
T_3	24,5	22,0	26,0	23,2	27,5
T_4	33,0	29,7	35,0	31,5	37,9
T_5	39,2	36,4	42,0	38,5	43,5

Таблиця порогових значень експертів

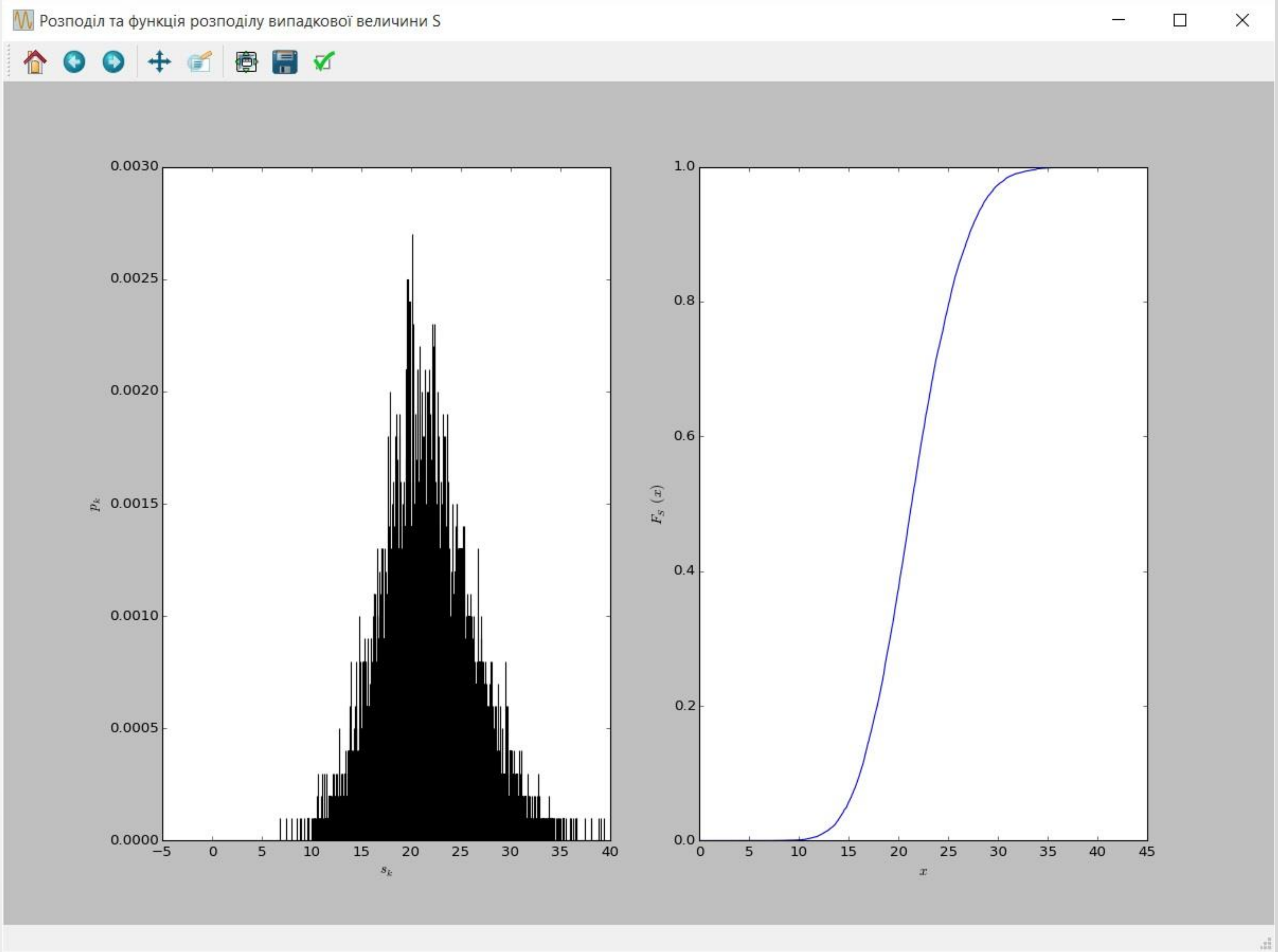
- Деякі типи помилок є корельованими.

Отримані результати

- За допомогою програмного продукту було отримано наступні імовірності отримання оцінок модельним студентом:

Оцінка	Імовірність отримання
A	0,0162
B	0,2721
C	0,5629
D	0,1471
E	0,0017
F	0,0

Розподіл оцінок для модельного студента



Розподіл та функція розподілу ВВ S

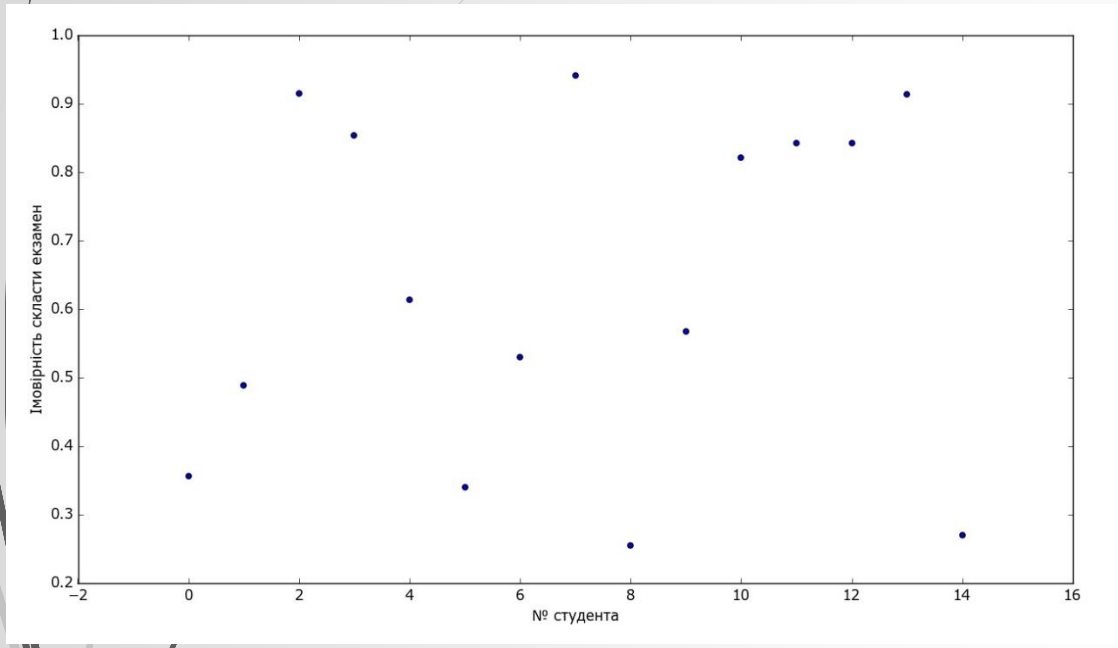
Задача знаходження розподілу кількості студентів, які потраплять на перездачу

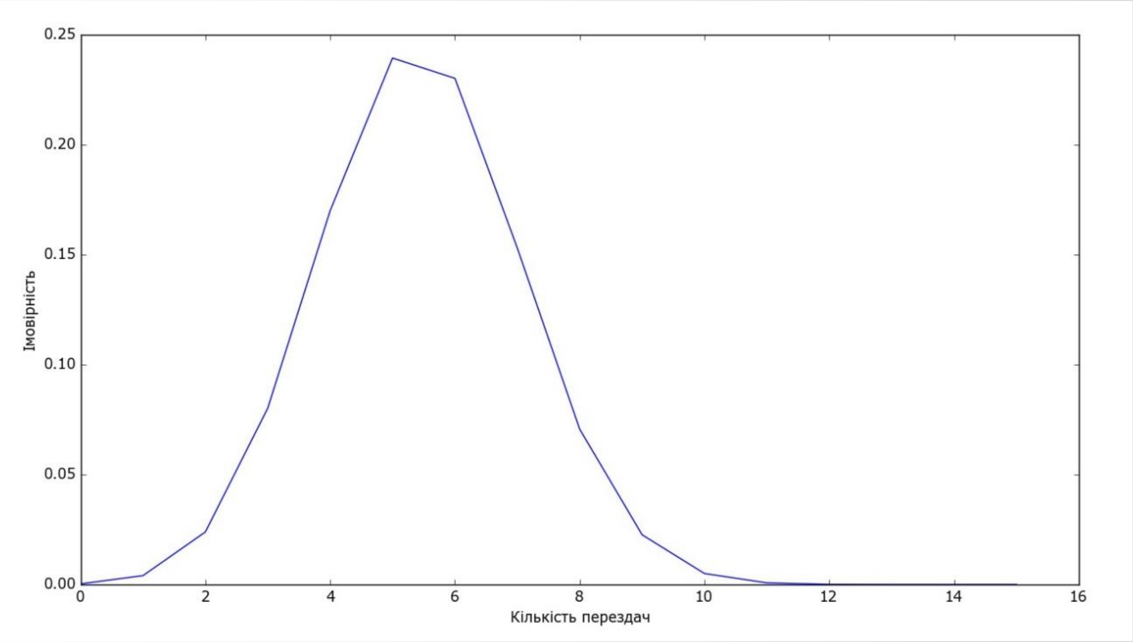
- Постановка задачі: Знаючи імовірності p_j успішного складання іспиту кожним j-м студентом, $j = \overline{1,n}$, знайти розподіл кількості студентів, які потраплять на перездачу.

- Задача зводиться до знаходження загальної кількості успіхів у n незалежних випробуваннях з різною імовірністю успіху, тобто до знаходження розподілу суми неоднаково розподілених бернуллівських ВВ.

- Така сума має пуассонівський біноміальний розподіл, існує рекурентна формула для знаходження імовірностей.

Приклад розв'язання задачі





Імовірності успішного складання іспиту для 15 студентів

Розподіл кількості студентів, що потраплять на перездачу

- Висновок: за таких умов на перездачу найімовірніше потрапить від 4 до 7 студентів.

Висновки

У результаті роботи:

- проведено математичну формалізацію задачі формування консолідованої експертної оцінки в припущенні, що якість виконання завдання визначається деякою ВВ S.
- проведено формалізацію ВВ S, яка враховує кількість і грубість допущених помилок, як зваженої суми пуассонівських ВВ;
- вивчено властивості цього розподілу у випадках незалежності у сукупності та залежності ВВ X_i . Це є важливим, оскільки розподіл виникає у багатьох застосуваннях, зокрема медичних та страхових, але є недостатньо вивченим;
- розглянуто існуючі та запропоновано власні підходи до точного обчислення та апроксимації функції розподілу ВВ S, знаходження точкових оцінок параметрів та довірчих інтервалів, серед них виявлено оптимальні;
- запропоновано альтернативні варіанти формалізації ВВ S;
- створено програмний продукт на основі моделі для автоматизації задання параметрів та виконання обчислень;
- протестовано створену модель на реальних даних.

Впровадження результатів дипломної роботи

- Побудована математична модель впроваджується в програмному продукті ТОВ «ВІМАС Технології» для вивчення іноземної мови іноземцями на основі технологій розпізнавання мовлення від SRI International.

- За результатами роботи подано статтю «Прогнозування результатів експертного оцінювання рівня знань іноземної мови за допомогою сум випадкових величин» до університетського наукового електронного збірника «Системні науки та кібернетика».

Дякую за увагу!