

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

**СТАТИСТИЧНІ МЕТОДИ
ТЕОРІЯ ОЦІНЮВАННЯ ТА СТАТИСТИЧНІ ГІПОТЕЗИ
ЛАБОРАТОРНИЙ ПРАКТИКУМ**

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для студентів, які навчаються
за спеціальністю 151 «Автоматизація та комп'ютерно-інтегровані технології»*

Київ
КПІ ім. Ігоря Сікорського
2018

Статистичні методи: Теорія оцінювання та статистичні гіпотези: Лабораторний практикум [Електронний ресурс] : навч. посіб. для студ. спеціальності 151 «Автоматизація та комп'ютерно-інтегровані технології» / КПІ ім. Ігоря Сікорського; уклад.: А. І. Жученко, Л. Д. Ярощук. – Електронні текстові дані (1 файл: 1,14 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2018. – 58 с.

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 10 від 21.06.2018 р.) за поданням Вченої ради Інженерно-хімічного факультету (протокол № 5 від 29.05.2018 р.)

Електронне мережне навчальне видання

СТАТИСТИЧНІ МЕТОДИ ТЕОРІЯ ОЦІНЮВАННЯ ТА СТАТИСТИЧНІ ГІПОТЕЗИ ЛАБОРАТОРНИЙ ПРАКТИКУМ

Укладачі: *Жученко А. І.*, завідувач кафедри «Автоматизація хімічних виробництв», доктор технічних наук, професор
Ярощук Людмила Дем'янівна, канд. техн. наук, доцент

Відповідальний редактор *Ковалюк Дмитро Олександрович*, канд. техн. наук, доцент

Рецензент: *Сокольський Олександр Леонідович*, к.т.н., доцент кафедри хімічного, полімерного та силікатного машинобудування інженерно-хімічного факультету КПІ ім. Ігоря Сікорського

Запропонований навчальний посібник містить матеріал для проведення лабораторного практикуму з методів дослідження технологічних об'єктів керування статистичними методами. Розглянуто дослідження законів розподілу випадкових величин; створення та дослідження графічних зображень емпіричних законів розподілу ймовірностей; визначення та дослідження оцінок параметрів законів розподілу випадкових величин; оцінювання тісноти зв'язку між випадковими величинами; ідентифікація технологічних об'єктів керування; визначення та відтворення автокореляційних функцій випадкових процесів. При виконанні лабораторних занять передбачено використання математичних процесорів *MathCAD* та *MS Excel*.

Призначений для студентів спеціальності «Автоматизація та комп'ютерно-інтегровані технології» усіх форм навчання.

© КПІ ім. Ігоря Сікорського, 2018

ЗМІСТ

	стор.
ВСТУП	4
Лабораторна робота №1 Дослідження законів розподілу випадкових величин.....	5
Лабораторна робота №2 Створення та дослідження графічних зображень емпіричних законів розподілу ймовірностей.....	13
Лабораторна робота №3 Дослідження оцінок параметрів законів розподілу випадкових величин.....	21
Лабораторна робота №4 Оцінювання тісноти зв'язку між випадковими величинами.....	30
Лабораторна робота №5 Ідентифікація технологічних об'єктів керування.....	35
Лабораторна робота №6 Дослідження автокореляційних функцій випадкових процесів	45
СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНИЙ ЛІТЕРАТУРИ	51
ДОДАТКИ	52

ВСТУП

Сучасний етап розвитку науки і техніки характеризується значними змінами методів наукових експериментів, аналізу та узагальнення отриманих результатів – все суттєвішою стає роль статистичних методів. Математико-статистичні дослідження стають необхідним інструментом для отримання більш глибоких знань про механізми явищ, що виникають. Значною мірою це стосується тих систем, структурна та динамічна складність яких робить неефективним чи взагалі неможливим використання аналітичних методів дослідження.

Навчальний посібник присвячений ознайомленню студентів з дослідженнями технологічних об'єктів керування статистичними методами.

При виконанні лабораторних робіт студенти вивчають методи визначення властивостей інформаційних потоків і створення математичних моделей об'єкта чи системи керування. Отримані моделі використовують в алгоритмах керування технологічними процесами для пошуку оптимальних режимів, прогнозування тощо.

Експериментально-статистичні методи фахівці активно застосовують завдяки значній кількості ефективних програмних продуктів, таких як *Statistica*, *MathCAD*, *MS Excel*, *MatLab* та ін.

У результаті виконання завдань до запропонованих лабораторних робіт студенти не тільки краще засвоюють теорію, але й оволодівають вміннями та досвідом використання статистичних функцій *MathCAD* та *MS Excel*.

Матеріал цього навчального посібника побудований таким чином, що його можна використати і при виконанні досліджень з іншими математичними пакетами.

Лабораторний робота №1

ДОСЛІДЖЕННЯ ЗАКОНІВ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН

Мета роботи – дослідити вплив параметрів законів розподілу ймовірностей на функції розподілу й щільності розподілу та на діапазон значень відповідних випадкових величин за допомогою спеціалізованої комп'ютерної системи *MathCAD*.

Стислі теоретичні відомості

Законом розподілу випадкової величини є будь-яка форма відповідності між значеннями цієї величини та ймовірностями появи цих значень. Закон розподілу може бути поданий як таблиця (табл.1.1) чи багатокутник розподілу (рис.1.1), а також за допомогою функцій розподілу ймовірностей $F(x)$ та щільності розподілу ймовірностей $f(x)$.

Таблиця 1.1 Таблиця розподілу ймовірностей

X	x_1	x_2	x_3	...	x_i	...	x_N
P	P_1	P_2	P_3	...	P_i	...	P_N

При упорядкуванні таблиці розподілу звичайно формують варіаційний ряд, тобто розташовують дані в порядку зростання. Таким чином, у наведеній таблиці дотримані співвідношення $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_i \leq x_N$.

Ряд розподілу можна описати і графічно. По осі абсцис відкладають значення випадкової величини, а по осі ординат - відповідні їм імовірності, точки з'єднують відрізками прямих ліній.

Отримана фігура називається *багатокутником розподілу*.

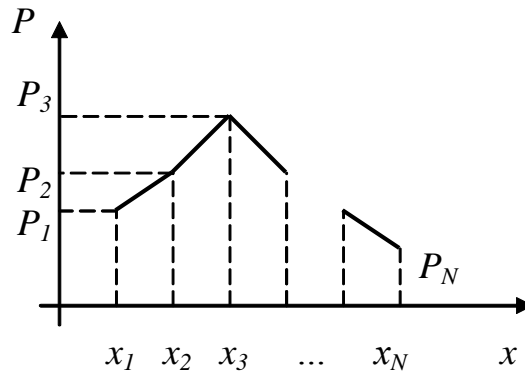


Рис.1.1. Багатокутник розподілу ймовірностей величини X

Таблиця та багатокутник розподілу використовуються для випадкових дискретних величин.

Найбільш загальною формою подання закону розподілу неперервних і дискретних випадкових величин є функція розподілу.

Функція розподілу, $F(x)$ - це ймовірність того, що випадкова величина X набуде значення, меншого за задане дійсне число x , тобто

$$F(x) = P(X < x) .$$

Імовірність попадання випадкової величини в інтервал $[\alpha, \beta]$ дорівнює різниці значень функції розподілу на кінцях цього інтервалу:

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha)$$

Функцією щільності розподілу ймовірностей, $f(x)$ є перша похідна від інтегральної функції $F(x)$, тобто $f(x) = F'(x)$. Вона існує тільки для випадкових неперервних величин.

Розглянемо декілька найбільш поширених стандартних законів розподілу.

Для дискретних величин:

- **біномний**, він має місце тоді, коли проводиться N незалежних дослідів, кожен з яких закінчується або успіхом, або невдачею (тобто може бути два взаємовиключаючих результати). Імовірність успіху дорівнює P , а невдачі – $q=1-P$. Значенням x випадкової величини X в такому випадку є кількість успіхів у серії з N дослідів, тобто X може мати тільки цілі додатні значення від 0 до N : $X = [0, N]$. Імовірність появи конкретного значення $X=x$ визначається так:

$$P_N(x) = P(X = x) = C_N^x \cdot P^x \cdot q^{N-x};$$

де
$$C_N^x = \frac{N!}{x!(N-x)!} .$$

Функція розподілу цього закону наступна

$$F(x) = \begin{cases} 0, & \text{при } x \leq 0; \\ \sum_{x_i \leq x} P_N(x_i), & \text{при } 0 < x \leq N; \\ 1, & \text{при } x > N; \end{cases}$$

- **пуассонівський**, цьому закону може підпорядковуватися випадкова величина, яка, як і в попередньому випадку, приймає цілі значення, а саме $X = [0, N]$, імовірність появи конкретного значення $X=x$ визначається так:

$$P_N(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}; \quad \lambda = N \cdot P .$$

Ця формула використовується тоді, коли N велике, а ймовірність появи події у окремому досліді (тобто P) незначна.

Для неперервних величин:

- **нормальний**, найбільш поширений серед величин, що характеризують технологічні процеси хімічної промисловості.

Функція щільності розподілу ймовірностей цього закону має вигляд

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2};$$

де μ , σ - параметри розподілу, а саме математичне сподівання та середнє квадратичне відхилення випадкової величини X ;

- **рівномірний**, функція щільності наступна

$$f(x) = \begin{cases} 0, & x \leq a; \quad x > b; \\ \frac{1}{b-a}, & a < x \leq b; \end{cases}$$

де a , b – параметри розподілу, а саме - граничні точки інтервала існування випадкової величини;

- **експоненціальний (показниковий)** має таку функцію щільності розподілу

$$f(x) = \begin{cases} 0, & x \leq 0; \\ \lambda e^{-\lambda x}, & x > 0, \end{cases}$$

де λ - параметр розподілу.

Mathcad має багату бібліотеку функцій роботи зі стандартними розподілами. Кожний розподіл може бути поданий трьома функціями – розподілу, щільності розподілу і функцією, оберненою до функції розподілу. Синтаксис кожної з них передбачає вживання на першому місці символів p , d , q відповідно, умовного позначення закону і в дужках певних аргументів (параметрів розподілу). В загальному випадку запис такий:

- d <ідентифікатор закону>(х,параметри закону) – повертає значення функції щільності розподілу для значення випадкової величини, що дорівнює x і має вказані параметри закону розподілу;
- p <ідентифікатор закону>(х,параметри закону) - повертає значення функції розподілу;
- q <ідентифікатор закону>(P,параметри закону) - повертає значення випадкової величини при заданій імовірності P і параметрах розподілу.

Наведемо синтаксис функцій розглянутих вище законів:

- **біномний:** $dbinom(x,N,P)$, $pbinom(x,N,P)$, $qbinom(P,N,r)$,

остання функція визначає кількість успіхів у N дослідах при умові, що імовірність успіхів не більше P , а r - імовірність успіху в одному досліді;

- **пуассонівський:** $dpois(x,\lambda)$, $ppois(x, \lambda)$, $qpois(P, \lambda)$;

- **нормальний:** $dnorm(x,\mu,\sigma)$, $pnorm(x,\mu,\sigma)$, $qnorm(P,\mu,\sigma)$;

- **рівномірний:** $dunif(x,a,b)$, $punif(x,a,b)$, $qunif(P,a,b)$;

- **експоненціальний:** $dexp(x, \lambda)$, $pexp(x, \lambda)$, $qexp(P, \lambda)$.

Порядок виконання роботи:

1. Для дискретних біномного та пуассонівського законів розподілу виконати дії у відповідності з рис. 1.2, а саме:

- відтворити документ *MathCAD*, зображений на рис.1.2 для біномного закону і створити самостійно схожий документ для пуассонівського закону розподілу;

- перевірити, чи дорівнює одиниці сума імовірностей всіх подій;

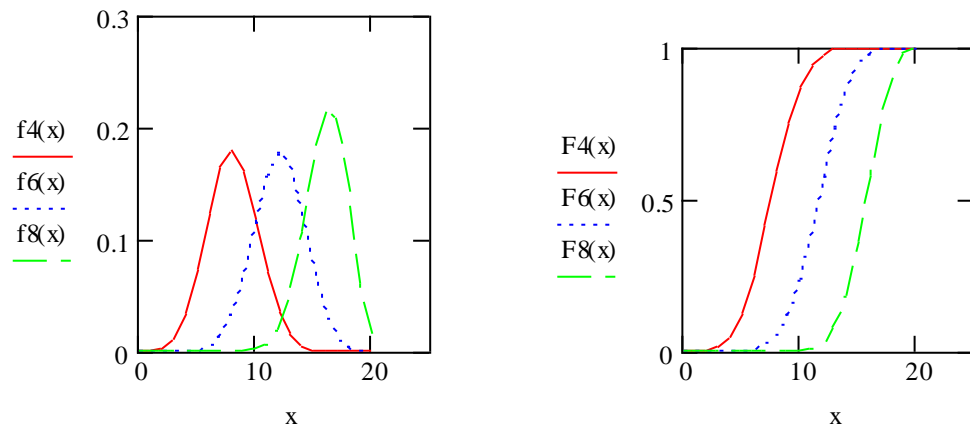
- для кожного розподілу визначити таке значення випадкової величини x , для якого $P(X=x)$ максимальна.

На рис.1.2 наведено приклад користування функціями *MathCAD* для біномного закону, коли значення випадкової величини (кількості успіхів) змінюється від 0 до 20 ($x:=0..20$), кількість дослідів $N=20$. У прикладі ймовірності успіху у одному досліді приймають послідовно такі значення: $P=0,4$; $P=0,6$; $P=0,8$. Зображення однотипних функцій подано на одному графіку.

$x := 0..20$

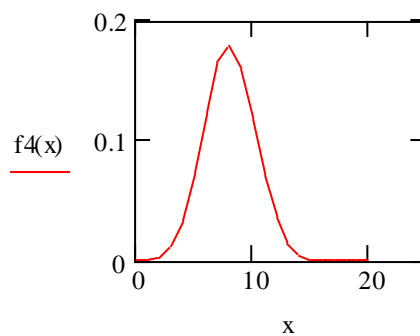
$f4(x) := \text{dbinom}(x, 20, 0.4)$ $f6(x) := \text{dbinom}(x, 20, 0.6)$ $f8(x) := \text{dbinom}(x, 20, 0.8)$

$F4(x) := \text{pbinom}(x, 20, 0.4)$ $F6(x) := \text{pbinom}(x, 20, 0.6)$ $F8(x) := \text{pbinom}(x, 20, 0.8)$



$$F4(9) - F4(4) = 0.704$$

$$\sum_{x=0}^{20} f4(x) = 1$$



$$f4_{\max} := 0.1781$$

$$x_{\max} := 8$$

Рис. 1.2. Приклад виконання розрахунків для біномного закону розподілу

На рис.1.2 наведено також розрахунок імовірності попадання значення випадкової величини у заданий інтервал, суму всіх значень імовірностей, а також найбільш імовірне значення випадкової величини для $f_4(x)$.

Для визначення найбільш імовірного значення треба, виділивши графік з $f_4(x)$, треба виконати команди **Формат-График-След...**

Після появи вікна **X-Y Trace**, покажчик миші слід встановити на максимумі $f_4(x)$.

2. Для неперервних розподілів дослідити вплив їхніх параметрів на вигляд функції розподілу $F(x)$ та функції щільності розподілу $f(x)$. Побудувати графіки $F(x)$ та $f(x)$ вказаних законів розподілу при наступних співвідношеннях параметрів (далі $N_{бр}$ – номер бригади):

- **нормального** : а) $\mu = N_{бр}$; • $\sigma = 0,1\mu$
 - $\sigma = \mu$;
- б) $\sigma = 1$; • $\mu = N_{бр}$
 - $\mu = 2N_{бр}$;
- **рівномірного** : а) $a = -N_{бр}$; $b = 2N_{бр}$;
- б) $a = N_{бр}$; $b = 4N_{бр}$;
- **експоненціального**: а) $\lambda = 0,5N_{бр}$; б) $\lambda = N_{бр}$; в) $\lambda = 1,5 N_{бр}$.

Зміст звіту

Звіт повинен вміщувати назву лабораторної роботи, мету дослідження, значення параметрів вказаних вище законів розподілу та відповідні графіки функцій $f(x)$ та $F(x)$, висновки.

Контрольні запитання

1. Що таке випадкова величина?
2. Що таке закон розподілу ймовірностей випадкової величини?
3. Які існують форми подання закон розподілу ймовірностей?
4. Наведіть приклади законів розподілу ймовірностей.
5. Назвіть параметри вищезазначених законів розподілу.
6. Які існують стандартні функції подання законів розподілу у MathCAD?
7. Як пов'язані параметри законів розподілу з діапазоном значень випадкових величин?

Лабораторна робота №2

СТВОРЕННЯ ТА ДОСЛІДЖЕННЯ ГРАФІЧНИХ ЗОБРАЖЕНЬ ЕМПІРИЧНИХ ЗАКОНІВ РОЗПОДІЛУ ЙМОВІРНОСТЕЙ

Мета роботи – дослідити вплив обсягів експериментальних даних на емпіричні закони розподілу ймовірностей випадкових величин.

Стислі теоретичні відомості

Отримані в експерименті значення $x_1, x_2, \dots, x_i, \dots, x_N$ випадкової величини X називаються варіантами (N - обсяг вибірки). Послідовність варіант, які записані у порядку зростання, формує так званий **варіаційний ряд**. Кількість спостережень, N_i кожного значення x_i називається **частотою**, а відношення:

$$W_i = N_i / N,$$

відносною частотою або **частістю**.

Наприклад, у результаті експериментальних досліджень отримана вибірка, що містить 50 варіант, які розподілені згідно з табл. 2.1. Треба визначити розподілення відносних частот.

Таблиця 2.1. Емпіричний розподіл величини X

X	5	8	14
N	10	15	25

Розв'язок. Обсяг вибірки становить: $N = \sum_{i=1}^3 N_i = 10 + 15 + 25 = 50$ варіант.

Відносні частоти для кожного значення X розраховують відповідно:

$$W_1 = 10/50 = 0,20; \quad W_2 = 15/50 = 0,30; \quad W_3 = 25/50 = 0,50; \quad W_1 + W_2 + W_3 = 1.$$

Емпіричним (статистичним) **розподілом** випадкової величини називається перелік варіант і відповідних їм частот або відносних частот, які спостерігались у вибірці експериментальних даних.

Емпіричний розподіл можна визначити також у вигляді послідовності інтервалів та суми частот, що опинилися в інтервалах. Для цього виконується групування даних спостережень у вигляді інтервального варіаційного ряду.

Використовуючи вибіркові дані, можна визначити емпіричну функцію розподілу, а також гістограму та полігон, які в певній мірі відображають функцію щільності розподілу. Позначимо через N_x кількість спостережень (варіант), при яких величини ознак (тобто значень X) були меншими за деяке x .

Емпіричною функцією розподілу називають функцію $F^*(x)$, яка для кожного значення x визначає відносну частоту події “ $X < x$ ”:

$$F^*(x) = N_x / N.$$

Різниця між емпіричною $F^*(x)$ та теоретичною $F(x)$ функціями розподілу полягає в тому, що $F(x)$ визначає імовірність події “ $X < x$ ”, а $F^*(x)$ характеризує відносну частоту цієї події. Згідно з теоремою Бернуллі, у великій кількості дослідів відносна частота наближається до імовірності $F(x)$ цієї події. Емпірична функція $F^*(x)$ має ті ж самі властивості, що і теоретична функція $F(x)$.

Побудуємо для прикладу емпіричну функцію розподілу за наведеними вище вибірковими даними.

Найменша варіанта дорівнює 5, отже $F(5)=0$.

Значення $x < 8$, тобто $x=5$, мали місце 10 разів, тому можна записати

$F^*(8)=10/50=0,2$. Значення $x < 14$, а саме $x=5$ та $x=8$ спостерігались $10+15=25$ разів, отже $F^*(14)=25/50=0,5$. Оскільки значення $x=14$ є найбільшим у вибірці, то можна визначити, що при $x > 14$ отримуємо $F^*=1$.

Таким чином,

$$F^*(x) = \begin{cases} 0 & \text{при } x \leq 5; \\ 0,2 & \text{при } 5 < x \leq 8; \\ 0,5 & \text{при } 8 < x \leq 14; \\ 1 & \text{при } x > 14. \end{cases}$$

Полігоном частот називають ламану лінію, відрізки якої з'єднують точки $(x_1, N_1), (x_2, N_2), \dots, (x_k, N_k)$, де k – кількість різних варіант. Для побудови полігону частот на осі абсцис відкладають варіанти x_i , а на осі ординат - відповідні частоти N_i , точки (x_i, N_i) з'єднують відрізками прямих ліній.

Полігоном відносних частот називають ламану лінію, відрізки якої з'єднують точки $(x_1, W_1), (x_2, W_2), \dots, (x_k, W_k)$. Для побудови полігону відносних частот на осі абсцис відкладають варіанти x_i , а на осі ординат - відповідні відносні частоти W_i .

Якщо результати вимірювань відносяться до неперервної випадкової величини X , то доцільно будувати гістограму.

Гістограмою частот називають східчасту фігуру, яка складається з прямокутників, основами яких є часткові інтервали довжиною Δ , а висоти дорівнюють відношенню N_i / Δ , яке називається щільністю частоти.

Гістограмою відносних частот називають східчасту фігуру, яка складається з прямокутників, основами яких є часткові інтервали довжиною Δ , а висоти дорівнюють W_i / Δ (щільність відносної частоти).

При побудові гістограми спочатку знаходять найменше і найбільше значення варіаційного ряду: x_{min} та x_{max} . Інтервал $[x_{min}, x_{max}]$, в якому містяться всі результати спостережень, розбивають на декілька часткових інтервалів, m довжиною Δ . Їх кількість визначається найчастіше за формулою Старджеса:

$$m = 1 + 3,322 \cdot \lg N;$$

отже довжину одиничного інтервалу визначають як

$$\Delta = (x_{max} - x_{min}) / (1 + 3,322 \cdot \lg N).$$

Початком першого інтервалу вважають величину $\delta_1 = x_{\min} - \Delta/2$, другого - $\delta_2 = \delta_1 + \Delta$, і-о - $\delta_i = \delta_{i-1} + \Delta$. Інтервали визначаються доти, доки початок наступного інтервалу не перевищить x_{\max} . Для кожного і-о часткового інтервалу знаходять N_i - кількість експериментальних даних, які потрапили до нього, а потім розраховують і відкладають по осі ординат щільність частоти (або відносної частоти в залежності від типу гістограми). Якщо інтервали мають різну довжину, то слід розглядати окремі Δ_i .

Mathcad має спеціальну функцію для побудови гістограм $hist(\delta, X)$. Результатом обчислень $hist(\delta, X)$ є вектор, кожний елемент якого дорівнює кількості вибіркового даних, значення яких належать відповідному інтервалу, тобто N_i . Параметр δ - вектор, в якому зберігаються граничні точки інтервалів. Розмірність вектора $hist(\delta, X)$ співпадає з розмірністю вектора δ і дорівнює кількості інтервалів. Визначивши необхідні складові $hist(\delta, X)$, можна отримати функцію $\varphi = hist(\delta, X)$.

Ця функція визначає кількість експериментальних значень у кожному інтервалі, її можна подати у вигляді гістограми чи полігона.. Для отримання конкретного виду зображення функції треба скористатися діалоговим вікном побудови графіків **Формат/Графік/Х-У Зависимость** (*Formating/Currently Selected/X-Y Plot*) і його опцією **След**. Для побудови гістограми вибирають установку **bar**, а для побудови полігону – **lines**.

Mathcad дозволяє генерувати випадкові величини, які підпорядковуються основним стандартним законам розподілу. Наведемо імена цих функцій для розглянутих законів: біномного $rbinom(N, k, P)$ – повертає вектор N випадкових чисел, які мають цей розподіл, k -натуральне число; пуассонівського $rpois(N, \lambda)$; нормального $rnorm(N, \mu, \sigma)$; рівномірного $runif(N, a, b)$; експоненціального $rexp(N, r)$.

Порядок виконання роботи

1. За допомогою *MathCad* побудувати графік емпіричної функції розподілу ймовірностей для даних наведених у табл. 2.2 (скористатись наведеним на рис.2.1 документом).

Таблиця 2.2. Результати експерименту

X	1	3	5	7	9
N	10	15	30	33	12

2. Побудувати гістограми та полігони за даними, які треба отримати в результаті генерування випадкові чисел (отримати вибірки для нормального, рівномірного та експоненціального розподілів). Приклад генерування даних та їхньої обробки наведено на рис. 2.2.

Використати наступні параметри розподілів для генерування даних:

- **нормальний закон** : $\mu = 10N_{бр}$; $\sigma = N_{бр}$;
- **рівномірний розподіл**: $a = -N_{бр}$; $b = 5N_{бр}$;
- **експоненціальний**: $\lambda = 0,5N_{бр}$.

Для кожного з розподілів сформувати вибірки при $N=20$; $N=200$; $N=200000$.

Зміст звіту

Звіт повинен вміщувати назву лабораторної роботи, мету дослідження, тексти документів *MathCAD* і графічні зображення результатів досліджень, висновки.

ORIGIN := 1

Створення матриці $A := \begin{pmatrix} -2 & 0 & 1 & 4 & 7 \\ 0.2 & 0.5 & 0.1 & 0.1 & 0.1 \end{pmatrix}$

Далі використовуємо інструмент програмування - Add Line
(панель інструментів "Математика" -> Програмування -> Add Line,
або Вид -> Панелі інструментов -> Программирование -> Add Line)

$$F(x) := \begin{cases} 0 & \text{if } -\infty < x < A_{1,1} \\ A_{2,1} & \text{if } A_{1,1} \leq x < A_{1,2} \\ ((A_{2,1} + A_{2,2})) & \text{if } A_{1,2} \leq x < A_{1,3} \\ ((A_{2,1} + A_{2,2} + A_{2,3})) & \text{if } A_{1,3} \leq x < A_{1,4} \\ ((A_{2,1} + A_{2,2} + A_{2,3} + A_{2,4})) & \text{if } A_{1,4} \leq x < A_{1,5} \\ 1 & \text{if } A_{1,5} \leq x < \infty \end{cases}$$

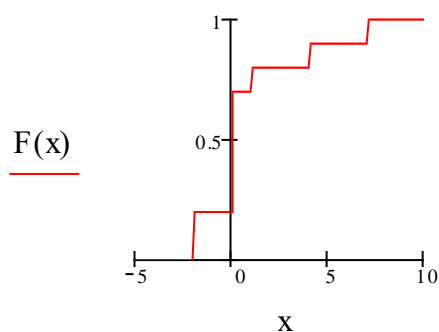


Рис.2.1. Документ для визначення емпіричної функції розподілу

ORIGIN := 1 N := 10 x := rnorm(10,1,0.1)

$$x^T =$$

	1	2	3	4	5	6	7	8	9	10
1	0.956	0.932	0.953	0.905	0.831	1.004	0.988	1.056	1.219	1.081

Xmax := max(x) Xmax = 1.219 Xmin := min(x) Xmin = 0.831
x := sort(x)

$$x^T =$$

	1	2	3	4	5	6	7	8	9	10
1	0.831	0.905	0.932	0.953	0.956	0.988	1.004	1.056	1.081	1.219

Rx := Xmax - Xmin Rx = 0.388

m := 1 + 3.322 · log(N) m = 4.322

m := floor(m) - ціла частина числа m , m = 4

$\Delta := \frac{Rx}{m}$ $\Delta = 0.097$ - довжина одного інтервалу

j := 1.. m + 2 $\delta_j := Xmin + (2 \cdot j - 3) \cdot \frac{\Delta}{2}$ f := hist(δ , x)

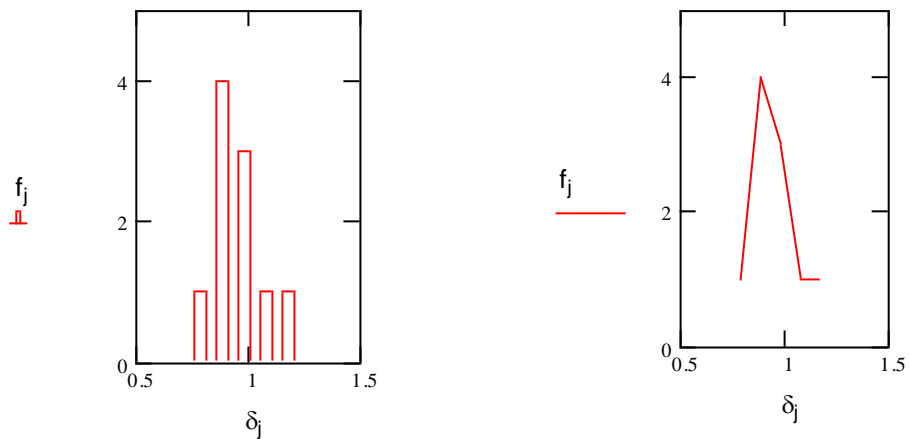


Рис.2.2.Документ для побудови гістограм і полігонів

Контрольні запитання

1. Що таке емпіричний закон розподілу ймовірностей?
2. Як побудувати емпіричну функцію розподілу ймовірностей дискретної випадкової величини?
3. Що таке полігон і гістограма частот (відносних частот)?
4. Як побудувати полігон та гістограму?
5. Як побудувати емпіричну функцію розподілу полігон та гістограму засобами *MathCAD*?

Лабораторна робота №3

ДОСЛІДЖЕННЯ ОЦІНОК ПАРАМЕТРІВ ЗАКОНІВ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН

Мета роботи – дослідити вплив обсягу вибірки та довірчої ймовірності на точкові та інтервальні оцінки параметрів законів розподілу випадкових величин.

Стислі теоретичні відомості

Закони розподілу цілком характеризують випадкові величини із імовірнісної точки зору. Кожний із них визначається одним чи декількома параметрами. Знання цих параметрів дозволяє отримати математичний вираз для функції розподілу, тобто повністю описати цей закон. Маючи, однак, лише вибіркові дані, можна обчислити тільки оцінки параметрів. Якщо дослідник визначився із видом закону розподілу, то він знає, які саме параметри треба розраховувати. Оцінки бувають точковими та інтервальними.

Точковою називають таку статистичну оцінку параметра, яка визначається одним числом. Значення такої оцінки θ^* невідомого параметра θ теоретичного розподілу, розраховується як функція від випадкових величин, що спостерігаються: $x_1, x_2, \dots, x_i, \dots, x_N$, ($x_i \in X$), у N незалежних дослідах.

Так, **нормальний закон**, якому відповідає наступна функція розподілу імовірностей випадкової величини X ,

$$F(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(z-\mu_x)^2}{2\sigma_x^2}} dz$$

має два параметри – математичне сподівання (μ_x) та дисперсію (σ_x^2).

Ці параметри є також найбільш поширеними числовими характеристиками випадкової величини з будь-яким законом розподілу ймовірностей.

Математичне сподівання визначає деяке середнє значення випадкової величини навкруг якого зосереджені всі імовірні її значення.

Незсуненою, обґрунтованою та ефективною оцінкою математичного сподівання генеральної сукупності, M_x є середнє арифметичне (M_a) результатів дослідження:

$$M_x = M_{\hat{a}} = \frac{x_1 + x_2 + \dots + x_N}{N}. \quad (3.1)$$

Параметри позначають звичайно грецькими літерами, а їх оцінки – латинськими.

Крім M_a середні значення випадкової величини бувають таких видів:

- **середнє квадратичне:** $M_{\hat{e}\hat{a}} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_N^2}{N}};$ (3.2)

- **середнє геометричне:** $M_{\hat{a}} = \sqrt[N]{x_1 x_2 \dots x_N};$ (3.3)

- **середнє гармонічне:** $M_{\hat{a}\hat{\delta}\hat{i}} = \frac{N}{1/x_1 + 1/x_2 + \dots + 1/x_N}.$ (3.4)

Узагальнюючою формулою для $M_{\text{кв}}$, M_a , $M_{\text{гарм}}$ є така:

$$M = \left[\frac{\sum_{i=1}^N x_i^k}{N} \right]^{1/k},$$

для $M_{\text{кв}}$: $k=2$; M_a : $k=1$; $M_{\text{Г}}$: $k=-1$.

Дисперсія σ_x^2 характеризує розсіювання випадкової величини навколо математичного сподівання. Її значення розраховують як математичне сподівання квадрата відхилення випадкової величини від математичного

сподівання:

$$\sigma_x^2 = \mu[X - \mu_x]^2.$$

Іноді використовують формулу

$$\sigma_x^2 = \mu_{x^2} - (\mu_x)^2.$$

Дисперсія має розмірність квадрата випадкової величини, тому часто вживають інший параметр - середнє квадратичне (стандартне) відхилення, σ_x :

$$\sigma_x = \sqrt{\sigma_x^2}.$$

Обґрунтованою та незсуненою є наступна точкова оцінка **дисперсії**:

$$S_x^2 = \frac{\sum_{i=1}^N (x_i - M_x)^2}{N - 1}. \quad (3.5)$$

Розглянемо інші поширені показники варіації:

розмах: $R = x_{\max} - x_{\min};$ (3.6)

де x_{\max}, x_{\min} – найбільше та найменше числа у вибірці;

середнє лінійне відхилення: $S_{\text{лин}} = \frac{\sum_{i=1}^N |x_i - M_x|}{N - 1};$ (3.7)

стандартне (середнє квадратичне) відхилення: $S_x = \sqrt{S_x^2};$ (3.8)

коефіцієнт варіації: $K_x = \frac{S_x}{M_x} 100\%.$ (3.9)

На рис. 3.1 наведено документ, у якому для розрахунку числових характеристик випадкових величин вжито стандартні функції *MathCAD*. Наведемо формули для деяких стандартних функцій у табл.3.1.

Робота з матрицями-рядками (чи матрицями-стовпцями)

$$A := (-4 \ 1 \ 1 \ 4 \ 7)$$

середнє $\text{mean}(A) = 1.8$

медіана $\text{median}(A) = 1$

дисперсія $\text{var}(A) = 13.36$

середнє квадратичне відхилення $\text{stdev}(A) = 3.655$

Робота з матрицями типу N*M

$$A := \begin{pmatrix} -2 & 0 & 1 & 4 & 7 \\ 2 & 0 & 4 & 3 & 0 \end{pmatrix}$$

$$\text{mean}(A) = 1.9 \quad \text{median}(A) = 1.5 \quad \text{var}(A) = 6.29 \quad \text{stdev}(A) = 2.508$$

Рис. 3.1. Стандартні статистичні функції *MathCAD* для розрахунку числових характеристик випадкових величин

Для випадку, коли випадкова величина підпорядковується **рівномірному закону розподілу** і існує на інтервалі $[0, \beta]$, не зсуненою та обґрунтованою є така оцінка β :

$$b = \frac{2}{N} \sum_1^N x_i. \quad (3.10)$$

Такими ж властивостями, але більш ефективною (тобто такою, що має меншу дисперсію) є наступна оцінка β :

$$b = \frac{N+1}{N} \max(x_1, x_2, \dots, x_N). \quad (3.11)$$

Таблиця 3.1. Таблиця відповідностей стандартних функцій *MathCAD*

Позначення	Формула
mean(x)	$\frac{1}{N} \sum_{i=1}^N x_i$
hmean(x)	$N \left(\sum_{i=1}^N 1/x_i \right)^{-1}$
gmean(x)	$\sqrt[N]{\prod_{i=1}^N x_i}$
var(x)	$\frac{1}{N} \sum_{i=1}^N (x_i - M_x)^2$
Var(x)	$\frac{1}{(N-1)} \sum_{i=1}^N (x_i - M_x)^2$
stdev(x)	$\sqrt{\text{var}(x)}$
Stdev(x)	$\sqrt{\text{Var}(x)}$

Нехай X – випадкова величина, яка підпорядковується експоненціальному закону розподілу і має наступну функцію щільності розподілу

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

Як видно з наведеної системи, цей закон визначається параметром λ .

Цей параметр може бути оцінений за вибірковими даними так

$$l = \frac{N}{\sum_1^N x_i} = \frac{1}{M_x}.$$

Точкові оцінки дають наближене значення параметру і самі є випадковими величинами. Їх дуже важко використовувати для прогнозування значень параметрів. Більш доцільно в останньому випадку вказувати діапазон, в якому може знаходитись значення параметру, тобто нижня і верхня границі його можливих значень: $M_x - \delta$, $M_x + \delta$.

Ці два числа і визначають так звану **інтервальну оцінку** параметра. Інтервальна оцінка визначається властивостями випадкової величини і заданою імовірністю знаходження оцінки у шуканому діапазоні. Діапазон значень, який має такі границі, називають довірчим інтервалом, а імовірність – **довірчою імовірністю** (P). Довжина діапазону дорівнює 2δ .

Спосіб визначення довірчого інтервалу для математичного сподівання нормально розподіленої випадкової величини залежить від того, чи відома заздалегідь її дисперсія. При відомій дисперсії і довірчій імовірності P довірчі границі розраховуються так:

$$M_x - \delta_\alpha \sqrt{\frac{\sigma_x^2}{N}}; M_x + \delta_\alpha \sqrt{\frac{\sigma_x^2}{N}},$$

де $\alpha=1-P$; δ_α - аргумент функції Лапласа $\Phi(\delta_\alpha)$, який визначається, виходячи з виразу

$$\Phi(\delta_\alpha) = P/2.$$

Параметр α називається рівнем значущості. Таблицю функції Лапласа наведено у табл.Д1 Додатку.

У тому разі, коли дисперсія апріорно не відома і можна розрахувати за вибірковими даними тільки її оцінку, довірчий діапазон визначають наступним

чином

$$\left[M_x - t_{\alpha, N-1} \sqrt{\frac{S_x^2}{N}}; M_x + t_{\alpha, N-1} \sqrt{\frac{S_x^2}{N}} \right],$$

де $t_{\alpha, N-1}$ - табличне значення критерію Стюдента, при рівні значущості α та $N-1$ степенях вільності для двосторонньої критичної області.

При заданому α визначають довірчу імовірність, як $P=1-0,5\alpha$ и по ній – значення критерію Стюдента.

Довірчий інтервал для дисперсії нормально розподіленої випадкової величини при відомому значенні її математичного сподівання μ і заданій довірчій імовірності P визначають наступним чином

$$\left[\frac{NS^2}{A}, \frac{NS^2}{B} \right].$$

Параметри A та B знаходять як корені рівнянь

$$X_{N,\alpha}^2(A) = 0,5\alpha \text{ і } X_{N,\alpha}^2(B) = 1 - 0,5\alpha,$$

де $X_{N,\alpha}^2$ - функція X^2 (хі-квадрат) – розподілу з N степенями вільності і рівні значущості α .

Якщо математичне сподівання апріорно не відоме, то довірчий інтервал для дисперсії визначається так

$$\left[\frac{((N-1)S^2)}{A}, \frac{(N-1)S^2}{B} \right].$$

Параметри A та B знаходять як корені рівнянь

$$X_{N-1,\alpha}^2(A) = 0,5\alpha \text{ і } X_{N-1,\alpha}^2(B) = 1 - 0,5\alpha,$$

де $X_{N-1,\alpha}^2$ - функція X^2 – розподілу з $N-1$ степенями вільності і рівні значущості α .

Порядок виконання роботи

1. Використовуючи функції *MathCAD*, генерувати вибірки обсягом $N=25$; $N=50$; $N=75$; $N=100$; $N=150$ наступних типів:

- **нормально розподілених випадкових величин** з такими параметрами:

$$\mu = N_{бр}, \sigma = 0,1N_{бр};$$

- **рівномірно розподілених випадкових величин** з параметром з діапазону $[0, \beta]$, $\beta = N_{бр}$;

- **випадкових величин**, які підпорядковуються **експоненціальному** закону розподілу з параметром $\lambda = N_{бр}$.

2. Для кожного N обчислити точкові оцінки параметрів вищеназваних законів розподілу, причому для рівномірного закону розподілу скористатися формулами (3.10) та (3.11).

3. Побудувати графіки, на яких зобразити дійсне значення параметру та залежність точкової оцінки від обсягу вибірки N .

4. Для п'яти вибірок нормально розподілених випадкових величин (див. п.1) розрахувати інтервальні оцінки математичного сподівання і дисперсії. Результати розрахунків занести у табл.3.2.

5. Побудувати графіки залежності довжини довірчих інтервалі

а) від обсягу вибірки N при вибраній довірчій імовірності ($P=0,95$);

б) від довірчої імовірності P при сталому обсязі вибірки ($N=50$).

6. Створити у *MS Excel* вибірку з 25 чисел, (використати дані, отримані за допомогою *MathCAD* у п.1). Розрахувати параметри цієї вибірки за (3.1-3.9), які вказано у теоретичних відомостях даної лабораторної роботи.

Таблиця 3.2. Результати розрахунків числових характеристик

Обсяг вибірки, N	Довірча імовірність, P			
	0,99	0,95	0,90	0,75
25				
50				
75				
100				
150				

Зміст звіту

Звіт повинен вміщувати назву лабораторної роботи, мету дослідження, формули для розрахунку точкових та інтервальних оцінок математичного сподівання та дисперсії, таблицю та графіки, отримані у результаті виконання пп.5,6, висновки.

Контрольні запитання

1. Які існують формули для визначення різних типів середніх значень випадкових величин та показників варіації?
2. Які відомі закони розподілу ймовірностей та їхні параметри?
3. Що таке оцінка параметру?
4. Що таке точкова оцінка параметру?
5. Що таке інтервальна оцінка параметру? Чим вона характеризується?
6. Які способи визначення параметрів надають *MS Excel* та *MathCAD* ?

Лабораторна робота №4

ОЦІНЮВАННЯ ТІСНОТИ ЗВ'ЯЗКУ МІЖ ВИПАДКОВИМИ ВЕЛИЧИНАМИ

Мета роботи – дослідити властивості показників статистичного зв'язку між випадковими величинами.

Стислі теоретичні відомості

Фактори, які описують стан різноманітних систем, можуть бути незалежними один від одного чи взаємопов'язаними. Взаємозв'язок розділяють на функціональний та стохастичний.

Перший має місце, коли система детермінована, тобто в ній відсутні неконтрольовані випадкові впливи. Якщо якийсь фактор залежить від іншого (інших), то ця залежність визначається на основі відомих законів природознавства. Математичні формулювання цих законів дозволяють розрахувати значення залежного фактора (Y) при кожному новому значенні фактора, що на нього впливає (X).

Реальні системи характеризуються, на жаль, наявністю випадкових впливів, значення яких не вимірюються. Зв'язок між факторами в цих системах відрізняється від функціонального. При стохастичній залежності при зміні однієї випадкової величини (X) змінюється закон розподілу іншої (Y). Система переходить не в один якийсь заздалегідь визначений стан, а в один із можливих станів. Такі системи також називаються стохастичними.

Окремим видом стохастичного зв'язку є кореляційний зв'язок. Вважається, що він має місце тоді, коли при зміні однієї величини (X) змінюється математичне сподівання іншої (Y).

Вираз

$$\mu[(X - \mu_x)(Y - \mu_y)]$$

називається **кореляційним моментом** чи **коваріацією** та позначається $\text{cov}(X, Y)$. Отже коваріація - це центральний змішаний момент другого порядку для величин X та Y . Якщо цей параметр не дорівнює нулю, то випадкові величини X та Y можна вважати залежними.

Вибіркове значення кореляційного моменту S_{xy}^2 розраховується так

$$S_{x,y}^2 = \frac{\sum_{i=1}^N (x_i - M_x)(y_i - M_y)}{N - 1}, \quad (4.1)$$

де M_x, M_y - вибіркові середні випадкових величин X та Y .

Тіснота розглянутих вище двох видів статистичного зв'язку визначається також кількома видами коефіцієнтів кореляції. Найбільш поширеним є коефіцієнт парної кореляції, r , який характеризує тісноту лінійного взаємозв'язку. Вибірковий коефіцієнт визначається за формулою

$$r = \frac{S_{x,y}^2}{S_x \cdot S_y}, \quad (4.2)$$

де S_x, S_y - середньоквадратичні (стандартні) відхилення випадкових величин X та Y .

Діапазон значень коефіцієнта r такий: $-1 \leq r \leq 1$. Чим більше модуль r , тим сильніший лінійний зв'язок між величинами. Значення "0" свідчить про відсутність лінійного зв'язку, а значення "1" чи "-1" вказують на функціональну залежність. Якщо $r > 0$, то X та Y одночасно збільшуються чи зменшуються, якщо $r < 0$, то при зростанні однієї величини інша зменшується. В загальному випадку для оцінки істотності лінійного статистичного зв'язку слід перевірити значущість r за критерієм Стьюдента. Для цього визначають розрахункове значення критерія за формулою

$$t = \frac{|r| \sqrt{N - 2}}{\sqrt{1 - r^2}}.$$

Потім перевіряють умову

$$t < t_{N-2,\alpha}, \quad (4.3)$$

де $t_{N-2,\alpha}$ - табличне значення критерія при $N-2$ ступенях вільності та рівні значущості α .

Ці розрахунки пов'язані з перевіркою статистичної гіпотези

$$H_0: \rho = 0$$

при альтернативній

$$H_1: \rho \neq 0,$$

де ρ - коефіцієнт парної кореляції, притаманний для генеральних сукупностей величин X та Y .

Вид альтернативної гіпотези свідчить про те, що розглядається двостороння критична область її прийняття. В цьому випадку при вибраному α треба або скористатися таблицею двостороннього критерія Стьюдента, або для таблиці одностороннього критерія прийняти $\alpha = \alpha/2$.

Якщо умова (4.3) виконується, то гіпотеза H_0 приймається, отже зв'язок між випадковими величинами слід визнати не істотним.

MathCAD має наступні стандартні функції для розрахунку вищезгаданих параметрів: $cvar(X,Y)$ – кореляційний момент, $corr(X,Y)$ – коефіцієнт парної кореляції (X,Y – експериментальні дані, задані матрицями-рядками чи матрицями-стовпцями).

Іноді виникають ситуації впливу якоїсь випадкової величини Z одночасно на декілька інших, наприклад, на X_1 та X_2 . Це призводить до похибки при оцінці тісноти зв'язку між X_1 чи X_2 та іншими величинами, наприклад, Y . В цьому випадку при дослідженні тісноти зв'язку між X_1 та Y , а також X_2 та Y дійсне його значення буде викривлене, причому чим сильніше впливає змінна Z на X_1 та X_2 , тим більшою буде похибка. Звільнитися від впливу Z можна двома шляхами - або проводити дослідження при сталому Z , або використовувати частковий коефіцієнт парної кореляції між X_1 та X_2 . Цей параметр визначається

за формулою

$$r_{yx_1, x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1 x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1 x_2}^2)}}, \quad (4.4)$$

де r_{yx_1, x_2} - частковий коефіцієнт кореляції між X_1 та Y при виключенні впливу X_2 .

Всі інші наведені в (4.2) параметри - коефіцієнти парної кореляції. Значення часткового коефіцієнта кореляції також належать діапазону $[-1, 1]$.

Порядок виконання роботи

1. Використати наведені в табл.4.1. експериментальні дані для розрахунку парних та часткових коефіцієнтів кореляції між всіма можливими парами змінних X_1 , X_2 та Y .
2. Розрахувати ті ж коефіцієнти при а) зміні знаку X_1 ; б) при одночасній зміні знаків X_1 , X_2 відносно п.1; в) при одночасній зміні знаків X_1 , X_2 та Y .

Зміст звіту

Звіт повинен вміщувати назву лабораторної роботи, мету дослідження, формули для розрахунку параметрів статистичного зв'язку та їхні значення, розраховані за різних умов у ході лабораторної роботи, висновки.

Контрольні запитання

1. Дати визначення функціонального, стохастичного та кореляційного типів зв'язків між величинами.
2. Як розрахувати кореляційний момент?
3. Що характеризує коефіцієнт парної кореляції? Як він розраховується?
4. Як визначити суттєвість кореляційного зв'язку?
5. Що характеризує коефіцієнт часткової кореляції? Як він розраховується?
6. Як визначити коефіцієнти парної та часткової кореляції у *MathCAD* та *MS Excel*?

Таблиця 4.1. Експериментальні дані для розрахунку коефіцієнтів кореляційного зв'язку

<i>N</i> п.п.	Витрата 1-го реагента (<i>X</i> ₁)	Витрата 2-го реагента (<i>X</i> ₂)	Концентрація продукту на виході (<i>Y</i>)
1	85,5	255	17,3
2	81,7	341	19,8
3	71,7	373	30,1
4	62,7	447	31,9
5	66,4	446	38,3
6	70,6	410	26,5
7	65,0	495	36,2
8	72,8	451	21,0
9	67,6	565	29,5
10	90,2	354	24,9
11	60,2	549	25,0
12	74,8	328	28,0
13	63,4	555	33,9
14	74,2	415	16,1
15	71,6	415	21,0
16	60,9	527	28,7
17	81,1	379	30,3
18	71,5	439	19,9
19	77,2	350	22,6
20	91,2	278	20,1

Лабораторна робота №5

ІДЕНТИФІКАЦІЯ ТЕХНОЛОГІЧНИХ ОБ'ЄКТІВ КЕРУВАННЯ

Мета роботи – дослідити особливості регресійного аналізу при ідентифікації технологічних об'єктів керування.

Стислі теоретичні відомості

Розглянуті раніше методи дисперсійного та кореляційного аналізу дозволяють визначити наявність зв'язку між випадковими величинами і оцінити силу цього зв'язку.

Наступним етапом повинно бути з'ясування форми зв'язку між випадковими величинами. Саме для отримання моделі об'єкта дослідження застосовується регресійний аналіз, який базується на експериментальних даних. Такий спосіб моделювання називається ідентифікацією.

Найчастіше зв'язок між факторами і вихідною змінною об'єкта у статистиці задається поліномом виду

$$y = \beta_0 + \sum_{i=1}^{N1} \beta_i x_i + \sum_{i=1}^{N1, N2} \beta_{ij} x_i x_j + \sum_{i=1}^{N3} \beta_{ii} x_i^2 + \dots$$

де β_0 , β_i , β_{ij} , β_{ii} - коефіцієнти, які характеризують відповідно середній рівень Y , лінійні ефекти, ефекти взаємодії між факторами та і квадратичні ефекти.

У залежності від виду рівняння регресії $y=f(x)$ розрізняють лінійну та нелінійну моделі. Вид функції $f(x)$ вибирається з огляду на фізико-хімічні особливості досліджуваного явища (процесу), а також на вид кореляційних полів. Величина найбільшого степеня фактора називається порядком моделі.

На етапі структурної ідентифікації створюють звичайно декілька конкуруючих моделей. Визначенню коефіцієнтів моделі присвячений етап параметричної ідентифікації.

Найпоширенішим способом розрахунку коефіцієнтів регресійних моделей є **метод найменших квадратів**.

Цей метод квадратів витікає з загального положення математичної статистики, згідно з яким за міру розсіювання завжди береться дисперсія.

Підберемо параметри рівняння так, щоб сума квадратів відхилень значень, що спостерігаються (y_i), від розрахункових (\hat{y}_i) була мінімальною, тобто

$$Q = \sum_{i=1}^N (\hat{y}_i - y_i)^2 \rightarrow \min.$$

Метод розрахунку параметрів, заснований на такій вимозі, називається методом найменших квадратів. Нехай

$$y = b_0 + b_1 x. \quad (5.1)$$

Розкриємо \hat{y}_i у наведеному вище функціоналі :

$$Q = \sum_{i=1}^N (b_0 + b_1 x_i - y_i)^2 \rightarrow \min.$$

Для відшукування мінімуму функціонала $Q(b_0, b_1)$, дорівняємо його частинні похідні по b_0 і по b_1 до нуля:

$$\begin{cases} \frac{\partial Q(b_0, b_1)}{\partial b_0} = 2 \sum_{i=1}^N (b_0 + b_1 x_i - y_i) = 0; \\ \frac{\partial Q(b_0, b_1)}{\partial b_1} = 2 \sum_{i=1}^N (b_0 + b_1 x_i - y_i) \cdot x_i = 0. \end{cases}$$

Розв'язавши цю систему, одержимо параметри b_0 і b_1 :

$$b_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{N \sum x_i^2 - (\sum x_i)^2}; \quad b_1 = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2}. \quad (5.2)$$

Метод оцінювання параметрів для багатофакторної моделі нічим не відрізняється від вищенаведеного метода оцінювання для однофакторної моделі, тільки розрахунки стають громіздкими.

Оскільки спостереження за кожною змінною дає не одне, а декілька її значень, то зручно записувати кожний такий ряд за допомогою вектора, а сукупність всіх незалежних змінних – за допомогою двовимірної матриці.

Введемо позначення, притаманні матричній формі запису:

$B=(b_j)$ - вектор оцінок параметрів (коефіцієнтів);

$Y=(y_i)$ - вектор значень залежної змінної, $i = \overline{1, N}$;

$X=(x_{ij})$ - матриця значень незалежних змінних розмірністю

$N \times M$, M - кількість незалежних змінних (факторів);

$\varepsilon=(\varepsilon_i)$ - вектор дійсних помилок;

$e=(e_i)$ - вектор оціночних помилок або залишків.

Запишемо модель з дійсними параметрами виду

$$\hat{y}_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots + \beta_M x_{M,i} + \varepsilon_i$$

у матричній формі так

$$Y = X \cdot \beta + \varepsilon ,$$

а для моделі з оцінками параметрів наступним чином

$$Y = X \cdot B + e .$$

Позначимо суму квадратів відхилень значень випадкової величини Y , розрахованих за моделлю, від експериментальних даних літерою J і запишемо її значення

$$\begin{aligned} J &= \sum e_i^2 = e^T e = (Y - X \cdot B)^T (Y - X \cdot B) = \\ &= Y^T Y - B^T X^T Y - Y^T X B + B^T X^T X B. \end{aligned}$$

Тут і далі “ T ” - позначення операції транспонування матриці.

Оскільки

$$B^T X^T Y = Y^T X B,$$

то

$$J = Y^T Y - 2B^T X^T Y + B^T X^T X B.$$

Продиференціюємо J по B :

$$\frac{\partial J}{\partial B} = -2X^T Y + 2(X^T X)B.$$

Оскільки $\frac{\partial J}{\partial B} = 0$, то можна записати

$$X^T Y = X^T X B,$$

звідки

$$B = (X^T X)^{-1} \cdot (X^T Y). \quad (5.3)$$

Розглянемо використані матриці детальніше.

Матрицю X запишемо у вигляді

$$X = \begin{vmatrix} 1 & x_{11} & x_{21} & \dots & x_{M1} \\ 1 & x_{12} & x_{22} & \dots & x_{M2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1N} & x_{2N} & \dots & x_{MN} \end{vmatrix}$$

(нагадаємо, що M - кількість факторів у моделі, N - кількість експериментів).

Таким чином, 1 - й рядок матриці X - це значення M вхідних змінних моделі у першому експерименті, N - й - у N -у експерименті.

Матриця Y - це вектор-стовпець:

$$Y = \begin{vmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{vmatrix}.$$

Тут y_i - значення вихідної змінної у i -у експерименті.

Зворотня матриця $(X^T X)^{-1}$ може існувати тільки тоді, коли початкова матриця $X^T X$ квадратна і її визначник не дорівнює нулю. Якщо ж визначник матриці дорівнює нулю, то вона вважається особливою. Матриця може виявитися особливою, якщо деякі з нормальних рівнянь є лінійними

комбінаціями інших рівнянь. У цьому випадку фактично має місце менша кількість рівнянь у порівнянні з кількістю невідомих параметрів.

У *MathCAD* для обчислення параметрів b_0 и b_1 (див. 5.2) використовуються функції відповідно $intersrpt(x,y)$ і $slope(x,y)$.

Розглянемо документ *MathCAD* (рис.5.1) з результатами ідентифікації об'єкта керування лінійною однофакторною моделлю типу (5.2), (дані взято з лабораторної роботи №4).

Документ *MathCAD* для визначення параметрів багатфакторної моделі відповідає формулі (5.3).

Microsoft Excel дозволяє розв'язати задачу декількома способами.

Розглянемо перший спосіб. Треба відкрити робочий документ і ввести числові дані у два стовпця. За допомогою **Мастера діаграмм** треба побудувати графік залежності Y від X , вибравши тип «точечная». Потім треба розташувати курсор миші на будь-якій з точок графіка і клацнути лівою кнопкою миші. Після цього зміниться колір усіх точок. Далі треба виконати команди меню **Диаграмма – Добавить линию тренда**. Відкриється діалогове вікно **Линия тренда**. Активізуємо вкладку **Тип** і виберемо тип лінії тренда – **линейная**, клацнувши по відповідному перемикачу. Крім цього зробимо активними два показчика: **Показывать уравнение на диаграмме** та **Показывать значение R^2 -квадрат на диаграмме** на вкладці **Параметры**. Натиснувши кнопку ОК, отримаємо кінцеве рішення на діаграмі.

Для використання іншої форми апроксимуючого рівняння в діалоговому вікні **Линия тренда** треба вибрати запропоновані варіанти.

ORIGIN:= 1

$X :=$	$\begin{pmatrix} 85.5 \\ 81.7 \\ 71.7 \\ 62.7 \\ 66.4 \\ 70.6 \\ 65 \\ 72.8 \\ 67.6 \\ 90.2 \\ 60.2 \\ 74.8 \\ 63.4 \\ 74.2 \\ 71.6 \\ 60.9 \\ 81.1 \\ 71.5 \\ 77.2 \\ 91.2 \end{pmatrix}$	$Y :=$	$\begin{pmatrix} 17.3 \\ 19.8 \\ 30.1 \\ 31.9 \\ 38.3 \\ 26.5 \\ 36.2 \\ 21.0 \\ 29.5 \\ 24.9 \\ 25.0 \\ 28.0 \\ 33.9 \\ 16.1 \\ 21.0 \\ 28.7 \\ 30.3 \\ 19.9 \\ 22.6 \\ 20.1 \end{pmatrix}$
--------	--	--------	--

$b0 := \text{intercept}(X, Y)$

$b1 := \text{slope}(X, Y)$

$b0 = 53.564$

$b1 = -0.377$

$i := 1..20$

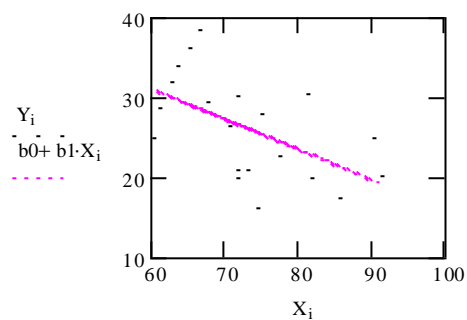


Рис.5.1. Документ *MathCAD* для визначення параметрів однофакторної моделі

На рис. 5.2 поданий документ *MS Excel*, у якому наведено розв'язок тієї ж задачі.

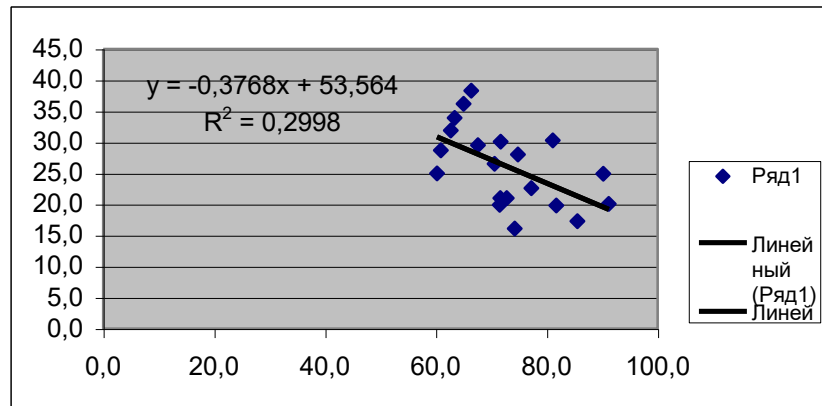


Рис. 5.2. Документ *MS Excel* з розв'язком задачі

Оцінка параметрів багатofакторної моделі в *MS Excel* виконується за допомогою стандартної функції ЛИНЕЙН.

Функція має такий синтаксис

ЛИНЕЙН(У;Х;константа;статистика),

де Y - вектор залежної змінної;

X - матриця незалежних змінних, де кожна змінна X_1, X_2, \dots, X_p є вектор стовпець, тобто

$$X = [X_1, X_2, \dots, X_p] = \begin{vmatrix} x_{11} & x_{21} & \dots & x_{p1} \\ x_{12} & x_{22} & \dots & x_{p2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{pn} \end{vmatrix}; \quad (5.4)$$

константа дорівнює 1, якщо в моделі є коефіцієнт b_0 , у противному разі константа дорівнює 0;

статистика – це логічне значення, що вказує, чи потрібно повернути додаткову інформацію стосовно регресійного рівняння.

Якщо статистика має значення 0, то функція ЛИНЕЙН повертає тільки рядок коефіцієнтів b_i . Якщо її значення дорівнює 1, то функція ЛИНЕЙН повертає додаткову регресійну статистику.

Для роботи з ЛИНЕЙН() треба виділити в робочому листі EXCEL блок комірок такого розміру: кількість стовпців дорівнює кількості параметрів моделі, а кількість рядків - п'ять.

Після цього викликати **Мастер функций** і вибрати у категорії **Статистические** функцію ЛИНЕЙН.

Виділити за допомогою миші діапазони комірок:

Y – вектор залежної змінної;

X - матрицю незалежних змінних

константа дорівнює 1;

статистика дорівнює 1.

Натиснути кнопку <Готово>, потім клавішу <F2>, а далі одночасно натиснути клавіші <Ctrl>+<Shift>+<Enter>.

Функція ЛИНЕЙН поверне додаткову регресійну статистику у вигляді табл.. 5.1.

Таблиця 5.1 Характеристики регресійної моделі

b_p	b_{p-1}	...	b_{p-2}	b_{p-1}	b_0
Sb_p	Sb_{p-1}	...	Sb_2	Sb_1	Sb_0
R^2	$S_{зал}$				
F	k				
$W_{регр}$	$W_{залт}$				

де $b_p; b_{p-1}; \dots; b_{p-2}; b_{p-1}; b_0$ – оцінки коефіцієнтів моделі;

$Sb_0 - Sb_p$ - стандартні похибки для оцінок коефіцієнтів;

R^2 – коефіцієнт детермінації $0 \leq R^2 \leq 1$;

$S_{\text{зал}} = \sqrt{S_{\text{зал}}^2}$, де $S_{\text{зал}}^2$ - залишкова дисперсія, яка визначається виразом

$$S_{\text{зал}}^2 = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N - (1 + M)}.$$

F – статистика, яка визначає, чи має взаємозв’язок між вихідною та факторами моделі випадковий характер. Для цього розраховується критерій Фішера за формулою

$$F = \frac{N - (1 + M)}{N - 1} \frac{\sum (y_i - M_y)^2}{\sum (y_i - \hat{y}_i)^2}.$$

Його значення показує, у скільки разів лінійна регресія краще описує взаємозв’язок між факторами та виходом, ніж середнє значення Y .

k – степінь вільності, який використовують для знаходження табличного значення критерію, $k = N - (1 + M)$;

$W_{\text{рег}}$ – регресійна сума квадратів, яку розраховують за формулою:

$$W_{\text{рег}} = \sum_{i=1}^N (\hat{y}_i - M_y)^2.$$

Залишкова сума квадратів розраховується так

$$W_{\text{зал}} = \sum_{i=1}^N (y_i - \hat{y}_i)^2.$$

Критичні значення статистичних критеріїв наведено у додатку (табл.Д2, Д3).

Порядок виконання роботи

1. Для експериментальних даних, наведених у лабораторній роботі №4, розрахувати параметри залежності $y=f(x_2)$ у *MathCAD* і *MS Excel*.
2. Отримати в *MathCAD* і *MS Excel* регресійну модель виду $y=f(x_1, x_2)$.
3. Визначити залишкові дисперсії кожної з моделей та множинні коефіцієнти кореляції.
4. Отримати графіки залишків, розрахувати їхні середні значення та дисперсії.

Зміст звіту

Звіт повинен вміщувати назву лабораторної роботи, мету дослідження, документи *MathCAD* та *MS Excel* з параметрами лінійної однофакторної моделі $y=f(x_2)$ і двофакторної моделі $y=f(x_1, x_2)$, графіки та параметри залишків, висновки.

Контрольні запитання

1. Які задачі розв'язує регресійний аналіз?
2. Що таке фактори, а що – параметри моделі?
3. Який метод застосовується для розрахунку параметрів регресійної моделі при нормальному розподілі випадкових величин?
4. Які способи проведення регресійного аналізу надають *MathCAD* та *MS Excel*?

ДОСЛІДЖЕННЯ АВТОКОРЕЛЯЦІЙНИХ ФУНКЦІЙ

Мета роботи – ознайомитися з типами автокореляційних функцій та дослідити алгоритми імітації числових послідовностей із заданими АКФ.

Стислі теоретичні відомості

Під час розробки імітаційної моделі об'єкта чи системи керування треба пам'ятати про динамічні властивості часових рядів. Використаємо поняття випадкової функції. *Випадковою функцією* називають функцію, яка в результаті досліду може прийняти той або інший конкретний вигляд, невідомий наперед. Вигляд, який приймає випадкова функція в результаті одного досліду, називають *реалізацією* цієї функції. Таким чином, числовий ряд можна сприймати як реалізацію випадкової функції.

У теорії ймовірностей основними числовими характеристиками однієї випадкової величини є *математичне сподівання* та *дисперсія*, а системи випадкових величин - *математичні сподівання* та *кореляційна матриця*.

Для випадкових функцій також уводять основні характеристики, аналогічні числовим характеристикам випадкових величин, і встановлюють правила дій з цими характеристиками. Але випадкові функції мають додаткову властивість – вони мають різне згасання залежності між значеннями елементів ряду при різних інтервалах вимірювання Δt . Ця властивість не характеризується ні математичним сподіванням, ні дисперсією.

Розглянемо для прикладу дві випадкові функції $X_1(t)$, $X_2(t)$, зображені сім'ями реалізацій на рис.6.1.

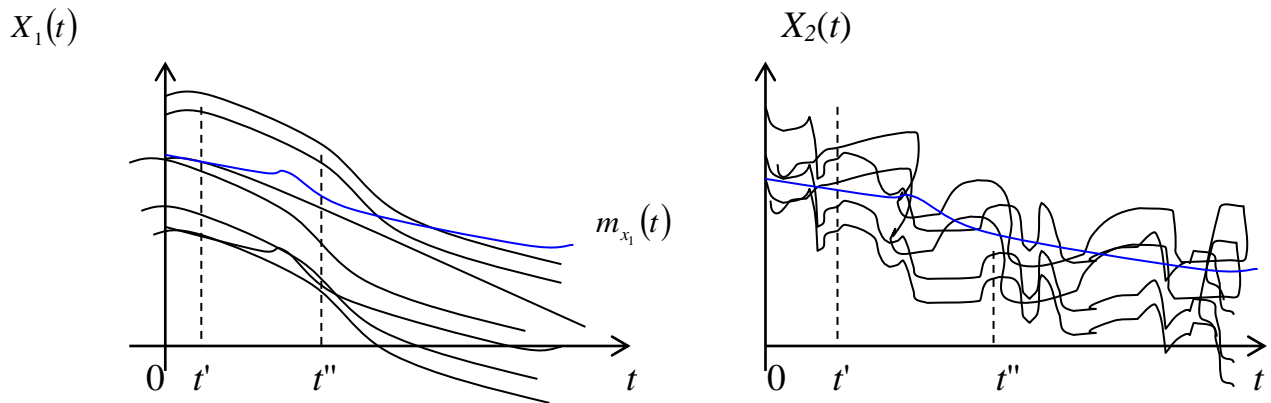


Рис.6.1. Реалізації випадкових функцій:
а) функції $X_1(t)$; б) функція $X_2(t)$

З рис.6.1 видно, що поведінка випадкових функцій у часі різна. Для її опису введена спеціальна характеристика, яку називають *кореляційною функцією* (автокореляційною функцією, АКФ). Таким чином, кореляційна функція характеризує ступінь залежності між перетинами випадкової функції, що належать до різних t .

Автокореляційна функція дає змогу оцінити інерційність інформаційного потоку, позначимо її $R_{xx}(\tau)$, тут τ – кількість інтервалів часу між вимірюваннями, ($t = \tau\Delta t$; $\tau = 0, 1, \dots, N - 1$).

Значення АКФ розраховують за вибірковими даними за формулою

$$R_{xx}(\tau) = \frac{1}{N - \tau - 1} \sum_{i=1}^{N-\tau} (x_i - M_x)(x_{i+\tau} - M_x).$$

Якщо інтервал між вимірюваннями нульовий, $\tau = 0$ (однакові числові ряди вимірювань), то

$$R_{xx}(0) = S_x^2.$$

Нормовану АКФ обчислюють за формулою

$$\bar{R}_{xx}(\tau) = \frac{R_{xx}(\tau)}{S_x^2}.$$

Отримавши графік вибіркової АКФ, дослідник може підібрати апроксимуючу функцію $\bar{R}_{xx,ap}(\tau)$. Для цього можна використати один з трьох поширених типів нормованої $\bar{R}_{xx,ap}(\tau)$:

$$\begin{aligned}\bar{R}_{xx,1}(\tau) &= \sigma_x^2 e^{-\alpha\tau}; \\ \bar{R}_{xx,2}(\tau) &= \sigma_x^2 e^{-\alpha\tau} \cos \beta\tau; \\ \bar{R}_{xx,3}(\tau) &= \sigma_x^2 [g_1 \bar{R}_{xx,1}(\tau) + g_2 \bar{R}_{xx,2}(\tau)],\end{aligned}\tag{6.1}$$

де α, β, g_1, g_2 – коефіцієнти апроксимації.

Похибку апроксимації розраховують за виразами

$$Q_1 = \frac{\sum_{\tau=0}^{\tau_{ap}} (\bar{R}_{xx}(\tau) - \bar{R}_{xx,ap}(\tau))^2}{N_{ap}},\tag{6.2}$$

або

$$Q_2 = \frac{\sum_{\tau=0}^{\tau_{ap}} |\bar{R}_{xx}(\tau) - \bar{R}_{xx,ap}(\tau)|}{N_{ap}},\tag{6.3}$$

де τ_{ap} – найбільша кількість інтервалів часу, вибрана для апроксимації, $N_{ap} = \tau_{ap} + 1$.

Апроксимуючу $\bar{R}_{xx}(\tau)$ далі будемо розглядати як задану АКФ.

Генерування числової послідовності з заданою АКФ базується на послідовностях нормально розподілених випадкових чисел і їхніх лінійних комбінаціях:

$$\begin{aligned}x(1h) &= c_1 u_1 + c_2 u_2 + \dots + c_N u_N; \\ x(2h) &= c_1 u_2 + c_2 u_3 + \dots + c_N u_{N+1}; \\ &\dots\dots\dots \\ x(Nh) &= c_1 u_N + c_2 u_{N+1} + \dots + c_N u_{2N-1},\end{aligned}\tag{6.4}$$

де u_i – елемент реалізації незалежної випадкової величини U із математичним сподіванням $\mu_u = 0$ та дисперсією $\sigma_u^2 = 1$.

Коефіцієнти c_k ($k = \overline{1, N}$) можна визначити з виразу

$$\overline{R}_{xx}(\tau) = (c_1 c_k + c_2 c_{k+1} + \dots + c_{N-k+1} c_N) S_x^2. \quad (6.5)$$

Розв'язок рівняння (6.5) не завжди можна отримати аналітичним шляхом, тому створено алгоритм генерування реалізацій випадкового процесу за наступними рекурентними виразами:

$$x_k = \sum_{i=0}^N c_i u_{k-i}; \quad k = 0, 1, \dots$$

$$x_k = \sum_{l=0}^L a_l u_{k-l} + \sum_{j=1}^m b_j x_{k-j}; \quad k = 0, 1, \dots$$
(6.6)

де a_l, b_j, c_i - параметри, які залежать від виду заданої кореляційної функції (див.4.1), яку треба відтворити.

У такому алгоритмі початкове значення x_0 можна прийняти нульовим. Це дещо викривлює початкову частину АКФ, але потім послідовність X стає стаціонарною.

Так, $\overline{R}_{xx,1}(\tau)$ випадкового процесу буде досягнута тоді, коли окремі значення ряду розраховувати за формулою:

$$x_k = a_0 u_k + b_1 x_{k-1}, \quad k = 0, 1, \dots, \quad (6.7)$$

де $a_0 = S_x \sqrt{1 - b_1^2}$; $b_1 = e^{-\alpha \cdot \tau}$; x_k, x_{k-1} - поточне та попереднє значення елементів ряду, що генерують.

АКФ $\overline{R}_{xx,2}(\tau)$ буде досягнута, якщо члени ряду розраховувати за виразом:

$$x_k = a_0 u_k + a_1 u_{k-1} + b_1 x_{k-1} + b_2 x_{k-2}, \quad k = 0, 1, \dots \quad (6.8)$$

Коефіцієнти a_0, a_1, b_1, b_2 визначають за формулами

$$c_0 = e^{-\alpha \tau} (e^{-2\alpha \tau} - 1) \cos \beta \tau; \quad c_1 = 1 - e^{-4\alpha \tau};$$

$$b_0 = \sqrt{(c_1 + \sqrt{c_1^2 - 4c_0^2})/2};$$

$$b_1 = 2e^{-\alpha \tau} \cos \beta \tau; \quad b_2 = e^{-2\alpha \tau};$$

$$a_0 = S_x b_0; \quad a_1 = S_x a_0 / b_0.$$

Для зменшення похибки наближення можна ввести підганяльні коефіцієнти δ_i , і тоді вирази (6.7, 6.8) приймуть вид:

$$x_k = a_0 u_k \delta_1 + b_1 x_{k-1} \delta_2, \quad k = 0, 1, \dots \quad (6.9)$$

$$x_k = a_0 u_k \delta_1 + a_1 u_{k-1} \delta_2 + b_1 x_{k-1} \delta_3 + b_2 x_{k-2} \delta_3, \quad k = 0, 1, \dots \quad (6.10)$$

Порядок виконання роботи

1. Отримати від викладача для досліджень стаціонарну числову послідовність $Y(t)$.
2. За допомогою *MathCAD* визначити її характеристики: M_y, S_y^2 .
3. Дослідити цю послідовність і підібрати з системи (6.1) таку АКФ, $R_{yy}(\tau)$, яка мала б найменшу похибку (за 6.2 або 6.3). Розглядати її як задану.
4. Згідно з M_y, S_y^2 та параметрами заданої $R_{yy}(\tau)$ за формулами (6.9, 6.10) згенерувати нову числову послідовність $X(t)$.
5. Дослідити цю послідовність, визначивши її вибіркові значення M_x, S_x^2 та підібравши для неї з (6.1) апроксимуючу АКФ, $R_{xx}(\tau)$.
6. Визначити точність відтворення $R_{yy}(\tau)$, порівнявши її з $R_{xx}(\tau)$ за (4.2) або (4.3).
7. Побудувати в одних координатах обидві АКФ - $R_{xx}(\tau)$ та $R_{yy}(\tau)$.

Зміст звіту

Назва, мета та порядок виконання роботи, контрольні запитання. Документи *MathCAD* з результатами виконання завдань. Висновки за результатами досліджень, зокрема про точність кожного зі способів імітації числових послідовностей із заданою АКФ.

Контрольні запитання

1. Яку функцію називають випадковою?
2. Назвіть основні характеристики випадкової функції.
3. Що характеризує автокореляційна функція випадкового процесу?
4. Що означає параметр τ у виразі АКФ?
5. Як розрахувати АКФ? Як визначити нормалізовану АКФ?
6. Наведіть найбільш поширені структури АКФ?
7. Наведіть показники для визначення точності апроксимації АКФ.

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Жученко, А. І. Спеціальні розділи математики для дослідження комп'ютерних систем [Текст]: навч. посіб / А. І. Жученко, Л. Д. Ярошук. – К.: Техніка, 1996. – 184с.
2. Жученко, А. І. Оцінювання параметрів та перевірка статистичних гіпотез. Теорія та практика роботи з MathCAD, MatLab, MS Excel [Текст]: навч. посіб / А. І. Жученко, Л. Д. Ярошук. – К.: НТУУ «КПІ», 2012. – 156 с.
3. Вентцель, Е.С. Теория вероятностей [Текст] / Е. С. Вентцель. – М.: Высш.шк., 2001. – 575 с.
4. Плис, А. И. Mathcad: математический практикум для экономистов и инженеров [Текст] / А. И. Плис, П. А. Сливина – М.: Финансы и статистика, 1999. – 656 с.
5. Дьяконов, В. Mathcad 2000 [Текст] / В. Дьяконов. – СПб.: Питер, 2001. – 592 с.
6. Кирьянов, Д. Самоучитель MathCAD 2001 [Текст] / Д. Кирьянов. – СПб.: bhv, 2001. – 544 с.
7. Блатнер, П. Использование Microsoft Excel 2000. Специальное издание [Текст] / П. Блатнер, Л. Ульрих. – М.: Издательский дом «Вильямс», 2000. – 1024 с.
8. Львовский, Е. Н. Статистические методы построения эмпирических формул [Текст] / Е. Н. Львовский. – М: Высш.шк., 1982. – 224 с.
9. Новикова, Л.В., Теорія ймовірностей і математична статистика [Текст] / Л. В. Новикова, Б. Д. Котляр, В. І. Бичков. – К.:Техніка, 1996. – 184 с.

ДОДАТКИ

Таблиця Д1. Значення функції Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$

X	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.0000	0.45	0.1736	0.90	0.3159	1.35	0.4115
0.01	0.0040	0.46	0.1772	0.91	0.3186	1.36	0.4131
0.02	0.0080	0.47	0.1808	0.92	0.3212	1.37	0.4147
0.03	0.0120	0.48	0.1844	0.93	0.3238	1.38	0.4162
0.04	0.0160	0.49	0.1879	0.94	0.3264	1.39	0.4177
0.05	0.0199	0.50	0.1915	0.95	0.3289	1.40	0.4192
0.06	0.0239	0.51	0.1950	0.96	0.3315	1.41	0.4207
0.07	0.0279	0.52	0.1985	0.97	0.3340	1.42	0.4222
0.08	0.0319	0.53	0.2019	0.98	0.3365	1.43	0.4236
0.09	0.0359	0.54	0.2054	0.99	0.3389	1.44	0.4251
0.10	0.0398	0.55	0.2088	1.00	0.3413	1.45	0.4265
0.11	0.0438	0.56	0.2123	1.01	0.3438	1.46	0.4279
0.12	0.0478	0.57	0.2157	1.02	0.3461	1.47	0.4292
0.13	0.0517	0.58	0.2190	1.03	0.3485	1.48	0.4306
0.14	0.0557	0.59	0.2224	1.04	0.3508	1.49	0.4319
0.15	0.0596	0.60	0.2257	1.05	0.3531	1.50	0.4332
0.16	0.0636	0.61	0.2291	1.06	0.3554	1.51	0.4345
0.17	0.0675	0.62	0.2324	1.07	0.3577	1.52	0.4357
0.18	0.0714	0.63	0.2357	1.08	0.3599	1.53	0.4370
0.19	0.0753	0.64	0.2389	1.09	0.3621	1.54	0.4382
0.20	0.0793	0.65	0.2422	1.10	0.3643	1.55	0.4394
0.21	0.0832	0.66	0.2454	1.11	0.3665	1.56	0.4406
0.22	0.0871	0.67	0.2486	1.12	0.3686	1.57	0.4418
0.23	0.0910	0.68	0.2517	1.13	0.3708	1.58	0.4429
0.24	0.0948	0.69	0.2549	1.14	0.3729	1.59	0.4441
0.25	0.0987	0.70	0.2580	1.15	0.3749	1.60	0.4452
0.26	0.1026	0.71	0.2611	1.16	0.3770	1.61	0.4463
0.27	0.1064	0.72	0.2642	1.17	0.3790	1.62	0.4474
0.28	0.1103	0.73	0.2673	1.18	0.3810	1.63	0.4484
0.29	0.1141	0.74	0.2703	1.19	0.3830	1.64	0.4495
0.30	0.1179	0.75	0.2734	1.20	0.3849	1.65	0.4505
0.31	0.1217	0.76	0.2764	1.21	0.3869	1.66	0.4515
0.32	0.1255	0.77	0.2794	1.22	0.3883	1.67	0.4525
0.33	0.1293	0.78	0.2823	1.23	0.3907	1.68	0.4535
0.34	0.1331	0.79	0.2852	1.24	0.3925	1.69	0.4545
0.35	0.1368	0.80	0.2881	1.25	0.3944	1.70	0.4554
0.36	0.1406	0.81	0.2910	1.26	0.3962	1.71	0.4564
0.37	0.1443	0.82	0.2939	1.27	0.3980	1.72	0.4573
0.38	0.1480	0.83	0.2967	1.28	0.3997	1.73	0.4582
0.39	0.1517	0.84	0.2995	1.29	0.4015	1.74	0.4591
0.40	0.1554	0.85	0.3023	1.30	0.4032	1.75	0.4599

Продовження табл.Д1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.41	0.1591	0.86	0.3051	1.31	0.4049	1.76	0.4608
0.42	0.1628	0.87	0.3078	1.32	0.4066	1.77	0.4616
0.43	0.1664	0.88	0.3106	1.33	0.4082	1.78	0.4625
0.44	0.1700	0.89	0.3133	1.34	0.4099	1.79	0.4633
1.80	0.4641	2.00	0.4772	2.40	0.4918	2.80	0.4974
1.81	0.4649	2.02	0.4783	2.42	0.4922	2.82	0.4976
1.82	0.4656	2.04	0.4793	2.44	0.4927	2.84	0.4977
1.83	0.4664	2.06	0.4803	2.46	0.4931	2.86	0.4979
1.84	0.4671	2.08	0.4812	2.48	0.4934	2.88	0.4980
1.85	0.4678	2.10	0.4821	2.50	0.4938	2.90	0.4981
1.86	0.4686	2.12	0.4830	2.52	0.4941	2.92	0.4982
1.87	0.4693	2.14	0.4838	2.51	0.4945	2.94	0.4984
1.88	0.4699	2.16	0.4846	2.56	0.4948	2.96	0.4985
1.89	0.4706	2.18	0.4854	2.58	0.4951	2.98	0.4986
1.90	0.4713	2.20	0.4861	2.60	0.4953	3.00	0.49865
1.91	0.4719	2.22	0.4868	2.62	0.4956	3.20	0.49931
1.92	0.4726	2.24	0.4875	2.64	0.4959	3.40	0.49966
1.93	0.4732	2.26	0.4881	2.66	0.4961	3.60	0.499841
1.94	0.4738	2.28	0.4887	2.68	0.4963	3.80	0.499928
1.95	0.4744	2.30	0.4893	2.70	0.4965	4.00	0.499968
1.96	0.4750	2.32	0.4898	2.72	0.4967	4.50	0.499997
1.97	0.4756	2.34	0.4904	2.74	0.4969	5.00	0.499997
1.98	0.4761	2.36	0.4909	2.76	0.4971		
1.99	0.4767	2.38	0.4913	2.78	0.4973		

Таблиця Д2. Критичні точки F-розподілу при $\alpha = 5\%$
 (v_1, v_2 – степені вільності дисперсій у чисельнику і у знаменнику
 критерія відповідно)

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.1172	9.1172	9.0135	8.9406	8.8868	8.8152	8.8103
4	7.7086	6.9443	6.5914	6.1881	6.3560	6.1694	6.0043	6.0410	5.9988
5	6.6679	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2066	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8378	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8626	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3808	3.5319	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3513	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3661
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0380	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9751	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2103	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2782	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0848	3.2317	2.8387	2.6060	2.4495	2.3359	2.2400	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2540	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6803	2.4472	2.2900	2.1750	2.0867	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

Продовження табл. Д2

$v_1 \backslash v_2$	10	12	15	20	24	30	40	60	120	120
1	241.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.385
3	10.128	9.5521	9.1172	9.1172	9.0135	8.9406	8.8868	8.8152	8.8103	8.8103
4	7.7086	6.9443	6.5914	6.1881	6.3560	6.1694	6.0043	6.0410	5.9988	5.9988
5	6.6679	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2066	4.1468	4.0990	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6767
8	5.3177	4.4590	4.0662	3.8378	3.6875	3.5806	3.5005	3.4381	3.3881	3.3881
9	5.1174	4.2565	3.8626	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4563
19	4.3808	3.5319	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.4227
20	4.3513	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3661	2.3661
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.3419
23	4.2793	3.4221	3.0380	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2821
26	4.2252	3.3690	2.9751	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2655
27	4.2103	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2782	2.2229	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.2107
40	4.0848	3.2317	2.8387	2.6060	2.4495	2.3359	2.2400	2.1802	2.1240	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2540	2.1665	2.0970	2.0401	2.0401
120	3.9201	3.0718	2.6803	2.4472	2.2900	2.1750	2.0867	2.0164	1.9588	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8799

Таблиця Д3. Критичні значення для t-критерія
 $P \{ \text{Студент } t \leq \text{табл. значення} \} = \gamma$

γ	0.75	0.90	0.95	0.975	0.99	0.995
1	1.0000	3.0777	6.3138	12.7062	31.8205	63.6567
2	0.8165	1.8856	2.9200	4.3027	6.9846	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0450	2.5705	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4249	1.8946	2.3646	2.9980	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7530	2.1314	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7237	2.0796	2.5177	2.8314
22	0.6853	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6843	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0513	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6613	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6516	1.3062	1.6835	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6803	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6339	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6320	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.5802	2.0154	2.4141	2.6923
45	0.6800	1.3006	1.6794	2.0141	2.4121	2.6896

Продовження табл. ДЗ

γ	0.75	0.90	0.95	0.975	0.99	0.995
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6346
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6600
50	0.6794	1.2987	1.6759	2.0086	2.4033	2.6778
51	0.6793	1.2984	1.6753	2.0076	2.4017	2.6757
52	0.6792	1.2980	1.6747	2.0066	2.4002	2.6737
53	0.6791	1.2977	1.6741	2.0057	2.3988	2.6718
54	0.6791	1.2974	1.6736	2.0049	2.3974	2.6700
55	0.6790	1.2971	1.6730	2.0040	2.3961	2.6682
56	0.6789	1.2969	1.6725	2.0032	2.3948	2.6665
57	0.6788	1.2966	1.6720	2.0025	2.3936	2.6649
58	0.6787	1.2963	1.6716	2.0017	2.3924	2.6633
59	0.6787	1.2961	1.6711	2.0010	2.3912	2.6618
60	0.6786	1.2958	1.6706	2.0003	2.3501	2.6603
61	0.6765	1.2955	1.6702	1.9996	2.3890	2.6589
62	0.6785	1.2954	1.6698	1.9990	2.3880	2.6575
63	0.6784	1.2951	1.6694	1.9983	2.3870	2.6561
64	0.6783	1.2949	1.6690	1.9977	2.3860	2.6549
65	0.6783	1.2947	1.6686	1.9971	2.3851	2.6536
66	0.6782	1.2945	1.6683	1.9966	2.3842	2.6524
67	0.6782	1.2943	1.6679	1.9560	2.3833	2.6512
68	0.6781	1.2941	1.6676	1.9955	2.3324	2.6501
69	0.6781	1.2939	1.6672	1.9949	2.3616	2.6490
70	0.6780	1.2938	1.6669	1.9944	2.3808	2.6479
71	0.6780	1.2936	1.6666	1.9939	2.3800	2.6469
72	0.6779	1.2934	1.6663	1.9935	2.3793	2.6459
73	0.6779	1.2933	1.6660	1.9930	2.3785	2.6449
74	0.6778	1.2931	1.6657	1.9925	2.3776	2.6439
75	0.6778	1.2929	1.6654	1.9921	2.3771	2.6430
76	0.6777	1.2928	1.6652	1.9917	2.3764	2.6421
77	0.6777	1.2926	1.6649	1.9913	2.3758	2.6412
78	0.6776	1.2925	1.6646	1.9908	2.3751	2.6403
79	0.6776	1.2924	1.6644	1.9905	2.3745	2.6395
80	0.6776	1.2922	1.6641	1.9901	2.3739	2.6387
81	0.6775	1.2921	1.0639	1.9897	2.3733	2.6379
82	0.6775	1.2920	1.6636	1.9893	2.3727	2.6371
83	0.6775	1.2918	1.6634	1.9890	2.3721	2.6364
84	0.6774	1.2917	1.6632	1.9836	2.3716	2.6356
85	0.6774	1.2916	1.6630	1.9883	2.3710	2.6349
86	0.6774	1.2915	1.6628	1.9879	2.3705	2.6342
87	0.6773	1.2914	1.6626	1.9876	2.3700	2.6335
88	0.6773	1.2912	1.6624	1.9873	2.3695	2.6329
89	0.6773	1.2911	1.6622	1.9870	2.3690	2.6322
90	0.6772	1.2910	1.6620	1.9867	2.3685	2.6316