МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КРЕМЕНЧУЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ МИХАЙЛА ОСТРОГРАДСЬКОГО
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЕЛЕКТРИЧНОЇ ІНЖЕНЕРІЇ
ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

МЕТОДИЧНІ ВКАЗІВКИ
ЩОДО ВИКОНАННЯ ПРАКТИЧНИХ І САМОСТІЙНОЇ РОБІТ З
НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«ІМОВІРНІСНО-СТАТИСТИЧНІ МЕТОДИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ»

ДЛЯ СТУДЕНТІВ ДЕННОЇ ФОРМИ НАВЧАННЯ ЗІ СПЕЦІАЛЬНОСТІ
123 – «КОМП'ЮТЕРНА ІНЖЕНЕРІЯ»
ОСВІТНЬО-ПРОФЕСІЙНОЇ ПРОГРАМИ
«КОМП'ЮТЕРНА ІНЖЕНЕРІЯ»
ОСВІТНЬОГО СТУПЕНЯ «БАКАЛАВР»
(ЧАСТИНА 2)

КРЕМЕНЧУК 2024

Методичні вказівки щодо виконання практичних і самостійної робіт студентів з навчальної дисципліни «Імовірнісно-статистичні методи інформаційних технологій» для студентів денної форми навчання для спеціальності 123 – «Комп’ютерна інженерія» освітньо-професійної програми «Комп’ютерна інженерія» освітнього ступеня «Бакалавр» (частина 2)

Укладач к. т. н., доц. В. М. Сидоренко
Рецензент д. т. н., проф. М. І. Гученко

Кафедра комп’ютерної інженерії та електроніки

Затверджено методичною радою Кременчуцького національного університету імені Михайла Остроградського
Протокол 5 від 23 02 2024 р.
Голова методичної ради проф. В. В. Костін

ЗМІСТ

ВСТУП

Навчальна дисципліна «Імовірнісно-статистичні методи інформаційних технологій» належить до циклу навчальних дисциплін математичної підготовки бакалаврської програми.

Мета практичних занять – узагальнити знання та навички, набуті під час опрацювання лекційного матеріалу та змістового модуля 1 «Теорія ймовірностей та ймовірнісні процеси» і навчитися розв’язувати відповідні задачі.

Методичні вказівки для студентів уміщують короткі теоретичні відомості, приклади розв’язування задач, вимоги щодо оформлення звіту індивідуальних завдань до тем «Випадкові величини», «Системи і функції випадкових величин», «Випадкові процеси», «Основи вибіркового методу».

Метою навчальної дисципліни є набуття студентами професійних компетенцій в галузі ймовірнісно-статистичних методів і підготовка студентів до ефективного їх використання в навчальному процесі, подальшій інженерній та науковій діяльності.

Завданням навчальної дисципліни є набуття знань закономірностей випадкових явищ і вміння використовувати ймовірнісно-статистичні методи для аналізу, моделювання та проектування апаратних і програмних складових комп’ютерних систем.

У результаті вивчення навчальної дисципліни студент повинен

отримати досвід з таких компетентностей.

ЗК 1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК 2. Здатність вчитися і оволодівати сучасними знаннями.

ЗК 3. Здатність застосовувати знання у практичних ситуаціях.

ЗК 7. Уміння виявляти, ставити та розв’язувати проблеми,

набути навички та уміння:

ПРН 1. Знати і розуміти наукові положення, на яких ґрунтується функціонування комп’ютерних засобів, систем та мереж.

ПРН 2. Мати навички проведення експериментів, збирання даних та моделювання в комп’ютерних системах.

ПРН 6. Уміти застосовувати знання для ідентифікації, формулювання і розв’язування технічних задач спеціальності, використовуючи методи, що є найбільш придатними для досягнення поставлених цілей.

ПРН 7. Уміти розв’язувати задачі аналізу та синтезу засобів, характерних для спеціальності.

ПРН 8. Уміти системно мислити та застосовувати творчі здібності до формування нових ідей.

ПРН 11. Уміти здійснювати пошук інформації в різних джерелах для розв’язання задач комп’ютерної інженерії.

ПРН 14. Уміти поєднувати теорію і практику, а також приймати рішення та виробляти стратегію діяльності для розв’язання завдань спеціальності з урахуванням загальнолюдських цінностей, суспільних, державних і виробничих інтересів.

ПРН 17. Спілкуватись усно та письмово з професійних питань українською мовою та однією з іноземних мов (англійською, німецькою, італійською, французькою, іспанською).

ПРН 18. Використовувати інформаційні технології для ефективного спілкування на професійному та соціальному рівнях.

ПРН 20. Усвідомлювати необхідність навчання впродовж усього життя з метою поглиблення набутих і здобуття нових фахових знань, удосконалення креативного мислення.

ПРН 21. Якісно виконувати роботу та досягати поставленої мети з дотриманням вимог професійної етики.

Згідно з вимогами освітньо-професійної програми студент повинен

знати:

– опис випадкових подій та їх аналіз;

– випадкові величини, системи і функції випадкових величин;

– математичний апарат опису і моделювання випадкових процесів та основи теорії СМО;

– основи вибіркового методу, точкові та інтервальні оцінки параметрів розподілу випадкових величин;

– основи теорії перевірки статистичних гіпотез;

– основи дисперсійного, кореляційного та регресійного аналізу;

уміти:

– обчислювати ймовірності випадкових величин з використанням комбінаторних формул;

– користуватися теоремами додавання та множення ймовірностей, формулами повної ймовірності та Байєса;

– обчислювати числові та функціональні характеристики випадкових величин;

– розв'язувати типові задачі теорії випадкових процесів, зокрема із застосування теорії СМО;

– реалізовувати задачі вибіркового методу та будувати моделі випадкових процесів для розв'язування задач статистичної обробки даних і побудови прогнозних моделей засобами спеціалізованої мови програмування R у середовищі RStudio;

– створювати проєкти з обробки та аналізу статистичних даних у середовищі RStudio за допомогою видавничої системи Quarto з використанням мов розмітки Markdown та LaTeX під контролем СКВ Git і розміщувати результати проєкту на GitHub.

1 ВИМОГИ ДО ОФОРМЛЕННЯ ЗВІТІВ ПРО ВИКОНАННЯ ЛАБОРАТОРНИХ РОБІТ

Звіт про виконання практичних робіт має бути написаний студентом від руки розбірливим почерком у зошиті в клітинку чи на комп'ютері шрифтом Times New Roman, розміром 14 пунктів на одному боці аркуша, через півтора інтервали. Відстань між попереднім текстом і заголовком має бути два інтервали, а відстань між заголовком і наступним текстом – у півтора рази більше, ніж міжрядковий проміжок звичайного тексту. Після заголовка на сторінці має бути хоча б один рядок тексту.

Зміст звіту

За підсумками кожної практичної роботи студент оформлює індивідуальний звіт, що містить:

– титульну сторінку;

– тему роботи;

– постановку завдання;

– розв'язання задачі згідно зі своїм варіантом;

– отримані результати;

– відповіді на контрольні питання.

2 ПЕРЕЛІК ПРАКТИЧНИХ РОБІТ

Практична робота № 5

Тема. Закони розподілу та числові характеристики дискретних випадкових величин. Закони розподілу та числові характеристики неперервних випадкових величин. Нормальний закон

Мета: набути практичних навичок у розв'язанні задач щодо знаходження законів розподілу та числових характеристик дискретних і неперервних випадкових величин, зокрема нормального закону, та розв'язання типових задач до цієї теми.

Короткі теоретичні відомості

Дискретні та неперервні випадкові величини

Визначення. Дискретною (перервною) (ДВВ) називають випадкову величину, яка набуває окремого ізольованого значення, із визначеними ймовірностями.

Визначення. Неперервною (НВВ) називають випадкову величину, яка може набувати всіх значень з деякого скінченного або нескінченного проміжку.

Визначення. Законом розподілу дискретної випадкової величини називають відповідність між можливими значеннями випадкової величини та їх імовірностями. Його можна задати у вигляді *таблиці*, *аналітично* (у вигляді формули) і *графічно*.

Приклад 5.1. Зіставимо дві сторони монети ДВВ відповідно: «Орел» – 1, «Решка» – 0. ДВВ X – випадіння нуля чи одиниці в одному киданні. Записати закон розподілу ДВВ у табличному вигляді.

Розв'язання.

X	0	1
p	$\frac{1}{2}$	$\frac{1}{2}$

$\sum p_i = p_1 + p_2 = 1.$

Біноміальний закон розподілу як приклад аналітично заданого закону

Визначення. Говорять, що випадкова величина X має *біноміальний* розподіл з параметрами n та p, де $0 \leq p \leq 1$, і пишуть $X \sim B_{n,p}$, якщо X набуває значень $0,1,\ldots,n$ з імовірністю $p(X = k) = C_n^k p^k (1-p)^{n-k}$.

Випадкова величина X з таким розподілом має сенс кількості успіхів у n випробуваннях схеми Бернуллі з імовірністю успіху p.

Математичне сподівання ДВВ

Нехай випадкова величина X може набувати тільки значень $x_1, x_2, \ldots, x_n$, імовірності яких відповідно дорівнюють $p_1, p_2, \ldots, p_n$.

Визначення. Математичним сподіванням (МС) ДВВ називають суму добутків усіх її можливих значень на їх ймовірності:

$$M(x) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n = \sum_{i=1}^{n} x_i p_i,$$

якщо випадкова величина X набуває *скінченної множини* можливих значень n та $M(x) = x_1 p_1 + x_2 p_2 + \ldots = \sum_{i=1}^{\infty} x_i p_i$, якщо випадкова величина X набуває з*численної множини* можливих значень.

Приклад 5.2. Знайти математичне сподівання ДВВ, заданою умовою задачі прикладу 1.1.

Розв'язання. Відповідно до визначення МС запишемо:

$$M(x) = \sum_{i=1}^{n} x_i p_i = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

Дисперсія ДВВ

Визначення. Дисперсією (розсіюванням) ДВВ називають математичне сподівання квадрата відхилення випадкової величини від її математичного сподівання:

$$D(x) = M[X - M(x)]^2 = \sum_{i=1}^{n} \left(X_i - M(x)\right)^2 p_i.$$

$$D(x) = M(x^2) - [M(x)]^2.$$

Визначення. Середнім квадратом відхилення (СКВ) випадкової величини X називають $\sigma(x) = \sqrt{D(x)}$.

Приклад 5.3. Знайти дисперсію ДВВ, заданою умовою задачі прикладу 5.1.

Розв'язання. Згідно з визначенням дисперсії ДВВ запишемо:

$$D(x) = M(x^2) - [M(x)]^2 = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} - \left[\frac{1}{2}\right]^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

Функція розподілу ймовірностей випадкової величини

Визначення. Функцією розподілу випадкової величини X називають функцію $F(x)$, яка визначає ймовірність того, що випадкова величина X у результаті випробувань набуде значення, менше x, тобто $F(x) = p(X < x)$.

Теорема. Функція розподілу $F(x)$ має такі властивості:

F1. Функція розподілу $F(x)$ неспадн**а**: $x_1 < x_2$, то $F(x_1) \leq F(x_2)$;

F2. Існують межі $\lim_{x \to -\infty} F(x) = 0$ та $\lim_{x \to \infty} F(x) = 1$;

F3. Функція розподілу $F(x)$ неперервна зліва: $F(x_0 - 0) = \lim_{x \to x_0} F(x) = F(x_0)$.

Функція розподілу є універсальною формою закону розподілу і може бути застосована як до дискретних, так і до неперервних випадкових величин.

Неперервні випадкові величини. Щільність імовірності та її ймовірнісний смисл

Визначення. Випадкова величина X має *абсолютно неперервний* розподіл, якщо існує невід'ємна функція $f(x)$ така, що для $\forall x \in R$ функція розподілу $F(x)$ подана у вигляді $F(x) = \int_{-\infty}^{x} f(t)\, dt$.

Функцію $f(x)$ називають *щільністю розподілу* випадкової величини X.

Теорема. Щільність розподілу має такі властивості:

f1. $f(x) \geq 0$ для $\forall x$;

f2. $\int_{-\infty}^{\infty} f(t)\, dt = 1$.

Приклад 5.4. Знайти функцію розподілу ДВВ, заданою умовою задачі прикладу 5.1 і побудувати її графік.

Розв'язання. $F(X) = p(x < X) = \begin{cases} 0, & \leq 0 \\ \frac{1}{2}, & 0 \leq x \leq 1. \\ 1, & x > 1 \end{cases}$

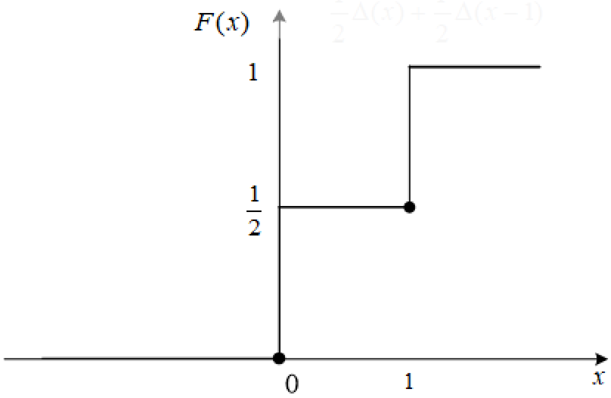

Рисунок 5.1 – Функція розподілу

Числові характеристики неперервної випадкової величини (НВВ). Математичне сподівання НВВ

Визначення. Математичним сподіванням НВВ X, можливі значення яких належать відрізку $[a, b]$, називають визначний інтеграл

$M(x) = \int_{\mathrm{a}}^{b} xf(x)\,dx$, якщо можливі значення належать усій осі Ox, то $M(x) = \int_{-\infty}^{\infty} xf(x)\,dx$.

Уважається, що невласний інтеграл збігається абсолютно, тобто існує інтеграл $\int_{-\infty}^{\infty} |x|\, f(x)dx$.

Дисперсія неперервної випадкової величини

Визначення. Дисперсією НВВ називають математичне сподівання квадрата її відхилення. Якщо можливі значення X належать відрізку $[a, b]$, то $D(x) = \int_{a}^{b} \big(x - M(x)\big)^2 f(x)dx$, якщо можливі значення X належать усій осі x, то $D(x) = \int_{-\infty}^{\infty} \big(x - M(x)\big)^2 f(x)dx$.

Визначення. СКВ НВВ визначається, як і для дисперсії ВВ:

$\sigma(x) = \sqrt{D(x)}$.

Приклад 5.5. Говорять, що X має рівномірний розподіл на відрізку $[a, b]$ і пишуть $X \sim U_{a,b}$, якщо $f(x) = \begin{cases} 0, x \le a \\ \frac{1}{b-a}, a < x \le b \\ 0, x > b \end{cases}$ (рис. 1.2),

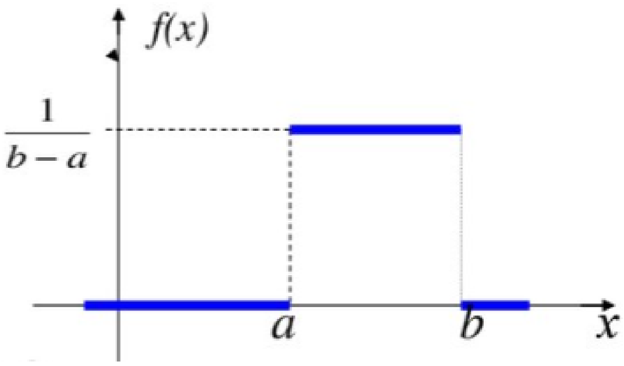

Рисунок 5.2 – Щільність рівномірного розподілу

Знайти функцію розподілу $F(x)$ та побудувати її графік. Знайти математичне сподівання $M(X)$ та дисперсію $D(X)$.

Розв'язання. Згідно з визначенням $F(x) = \int_{-\infty}^{x} f(t)\,dt$. Якщо $x \le a$, то $f(x) = 0$, отже, $F(x) = 0$. Якщо $a < x \le b$, то $f(x) = \frac{1}{b-a}$, отже,

$$F(x) = \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{a} 0dx + \int_{a}^{x} \frac{1}{b-a} dx = \frac{x-a}{b-a}.$$

Якщо $x > b$, то

$$F(x) = \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{a} 0dx + \int_{a}^{b} \frac{1}{b-a} dx + \int_{b}^{\infty} 0dx = \frac{b-a}{b-a} = 1.$$

Отже, шукана функція розподілу (рис. 1.3):

$$F(x) = P(X < x) = \begin{cases} 0, x \le a \\ \frac{x-a}{b-a}, a < x \le b. \\ 1, x > b \end{cases}$$

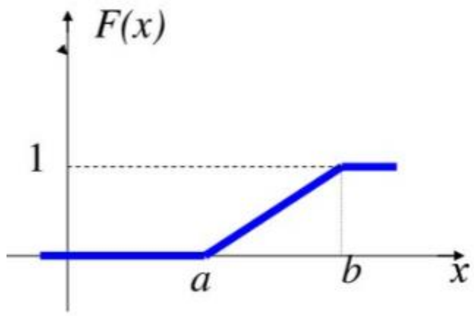

Рисунок 5.3 – Функція розподілу рівномірного розподілу

Нормальний розподіл

Визначення. Говорять, що X має нормальний розподіл з параметрами μ і σ^2, де $\mu \in R$, $\sigma > 0$, і пишуть $X \sim N(\mu; \sigma^2)$, якщо X має таку щільність розподілу (рис. 1.4):

$f(x) = \frac{1}{2 \cdot \sigma \sqrt{2\pi}} \mathrm{e}^{\frac{-(x-\mu)^2}{2\sigma^2}}$ для $\forall x \in R$, $M(x) = \mu$, $D(x) = \sigma^2$.

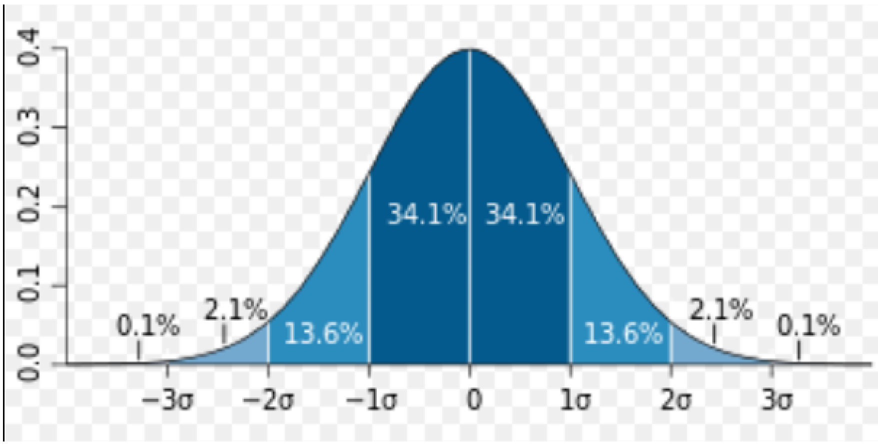

Рисунок 5.4 – Крива нормального розподілу $X \sim N(0; 1)$, де кожна вертикальна смуга має ширину, що дорівнює одному стандартному відхиленню. (За даними Вікіпедії)

Властивості нормального розподілу

1. $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) + \frac{1}{2} = \frac{1}{2} + \int_0^x e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt$
2. $p(x_1 < X < x_2) = \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right)$
3. $\Phi(\infty) = 0{,}5$
4. $\Phi(-x) = -\Phi(x)$
5. Імовірність заданого відхилення від середнього на величину δ:

$$P(|X-\mu| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right)$$

Правило «трьох сигм»

$P(|X-\mu| < 3\sigma) = 2\Phi(3) = 0{,}9973.$

$P(|X-\mu| > 3\sigma) = 1 - 0{,}9973 = 0{,}0027.$

Приклад 5.6. НВВ X має нормальний розподіл з параметрами $\mu = 3$ та $\sigma = 1$. Функція щільності нормального розподілу $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}$. Обчислити ймовірність події $1 \leq X \leq 2$ та ймовірність ($P(|x - \mu| \leq \delta)$. того, що значення випадкової величини відхилиться від математичного сподівання $\mu = 3$ на величину, що не перевищує $\delta = 0{,}01$.

Розв'язання. Згідно з властивістю 2, наведеною вище, можна записати $p(1 \leq X \leq 2) = \Phi\left(\frac{2-3}{1}\right) - \Phi\left(\frac{1-3}{1}\right) = \Phi(-1) - \Phi(-2) = \Phi(2) - \Phi(1) = 0{,}4772 - 0{,}3413 = 0{,}1359$. Тобто ≈ 14 шансів на 100, що значення ВВ опиниться у цих межах.

Для знаходження ймовірності $P(|x - 3| \leq 0{,}01)$, скористаємося властивістю 5: $P(|X - 3| < 0{,}01) = 2\Phi\left(\frac{0{,}01}{3}\right) \approx 2 \cdot 0{,}000 = 0{,}000$.

Задачі для самостійного розв'язання

Виконати індивідуальне завдання. Завдання полягає у розв'язанні п'яти задач, які потрібно вибрати зі списку, наведеного нижче. Правило вибору номерів таке: $n, n+1, n+2, n+3, n+4$, де n – номер студента у списку групи. У разі, якщо було досягнуто кінця списку задач, потрібно циклічно повернутися до його початку.

Завдання

1. У мішень виконується 4 незалежні постріли з імовірністю влучення внаслідок кожного пострілу $p = 0{,}8$. Необхідно: 1) знайти закон розподілу ДВВ X, що дорівнює кількості влучень у мішень; 2) виразити функцію розподілу та функцію щільності розподілу ДВВ за допомогою функції Хевісайда та δ-функції Дірака; 3) побудувати графіки функцій розподілу та щільності розподілу; 4) знайти ймовірності подій $1 \leq x \leq 3$ та $x > 3$; 5) побудувати багатокутник розподілу; 6) знайти математичне сподівання, дисперсію, середнє квадратичне відхилення, теоретичні початкові та центральні моменти 3-го та 4-го порядку; 7) знайти асиметрію та ексцес.

2. З імовірністю влучення внаслідок одного пострілу $p = 0.7$ стрілок стріляє у мішень до першого влучення, але може виконати не більше, ніж 4 постріли. ДВВ X – кількість промахів. Необхідно: 1) знайти закон розподілу ДВВ; 2) виразити функцію розподілу та функцію щільності розподілу ДВВ за допомогою функції Хевісайда та δ-функції Дірака; 3) побудувати графіки функцій розподілу та щільності розподілу; 4) знайти асиметрію та ексцес.

3. Двічі кинута гральна кістка. ДВВ X – різниця між кількістю очок унаслідок першого кидання та кількістю очок унаслідок другого кидання. Необхідно: 1) знайти закон розподілу ДВВ; 2) побудувати графік функції щільності розподілу ДВВ; 3) знайти ймовірність події.

4. В урні 7 кульок, із яких 4 білі, а інші – чорні. З цієї урни навмання беруть 3 кульки. ДВВ X – кількість білих кульок. Необхідно: 1) знайти закон розподілу ДВВ; 2) виразити функцію розподілу та функцію щільності розподілу ДВВ за допомогою функції Хевісайда та δ-функції Дірака; 3) побудувати графіки функцій розподілу та щільності розподілу; 4) знайти ймовірність події $x \geq 1$; 5) побудувати багатокутник розподілу; 6) знайти математичне сподівання, дисперсію, середнє квадратичне відхилення, теоретичні початкові та центральні моменти 3-го та 4-го порядків; 7) знайти асиметрію та ексцес.

5. Завод відправив на базу 500 цілих деталей. Імовірність зіпсування кожної деталі в дорозі $p = 0{,}002$. Знайти закон розподілу ДВВ X, що дорівнює кількості зіпсованих деталей, і знайти ймовірності подій:

– пошкоджено менше, ніж 3 деталі;
– пошкоджено більше, ніж 2 деталі;
– пошкоджено хоча б одну деталь.

6. Два стрілки роблять по одному пострілу в одну мішень. Імовірність влучення для першого стрілка внаслідок одного пострілу $p_1 = 0{,}5$, для другого – $p_2 = 0{,}4$. ДВВ X – кількість влучень у мішень. Необхідно: 1) знайти закон розподілу ДВВ X, що дорівнює кількості влучень у мішень; 2) виразити функцію

розподілу та функцію щільності розподілу ДВВ за допомогою функції Хевісайда та δ-функції Дірака; 3) побудувати графіки функцій розподілу та щільності розподілу; 4) знайти ймовірності подій $1 \le x \le 3$ та $x > 3$; 5) побудувати багатокутник розподілу; 6) знайти математичне сподівання, дисперсію, середнє квадратичне відхилення, теоретичні початкові та центральні моменти 3-го та 4-го порядків; 7) знайти асиметрію та ексцес.

7. НВВ X має рівномірний розподіл з параметрами a, b. Функція щільності рівномірного розподілу $f(x) = \frac{1}{b-a}, a \le x \le b$. Вивести формулу функції рівномірного розподілу $F(x)$, формулу для математичного сподівання $M(x)$, дисперсії $D(x)$, асиметрії As, ексцесу Ek, імовірності події $\alpha \le X \le b$.

8. НВВ X має експоненціальний розподіл з параметром λ. Функція щільності експоненціального розподілу $f(x) = \lambda e^{-\lambda x}, x \ge 0$. Вивести формулу функції рівномірного розподілу $F(x)$, формулу для математичного сподівання $M(x)$, дисперсії $D(x)$, імовірності події $\alpha \le X \le b$.

9. НВВ X має розподіл Коші. Функція щільності розподілу Коші задана у вигляді $f(x) = \frac{c}{1+x^2}$, де c – деяка константа. Знайти константу c, функцію розподілу Коші $F(x)$ та ймовірність події $-1 \le X \le 1$.

10. НВВ X задана функцією щільності розподілу:

$$f(x) = \begin{cases} c \cdot cos\, x, -\frac{\pi}{2} \le x \le \frac{\pi}{2}, \\ 0, |x| > \frac{\pi}{2} \end{cases}, \text{ де } c \text{ – деяка константа.}$$

Знайти константу c, функцію розподілу $F(x)$, імовірність події $|X| \le \frac{\pi}{4}$.

11. Часовий інтервал між надходженнями пакетів даних у комп’ютерній мережі зі швидкістю передавання даних 10 *Мбіт/сек* підприємства має експоненціальний розподіл з параметром $\lambda = 15$ *мкс*. Знайти:

– середню довжину інтервалу;

– дисперсію довжини інтервалу;

– СКВ довжини інтервалу;

– імовірність того, що часовий інтервал між надходженнями пакетів перевищить 20 $мкс$;

– імовірність того, що часовий інтервал між надходженнями пакетів буде у межах $10 < \text{X} < 15$ $мкс$.

12. Параметри генератора псевдовипадкових чисел $random$, що входить до інтегрованого середовища $Turbo\ Pascal$, за замовчанням дорівнюють $a = 0$, $b = 1$. Знайти:

– математичне сподівання;

– дисперсію;

– середнє квадратичне відхилення;

– імовірність того, що $X > 0.5$.

13. НВВ X має нормальний розподіл з параметрами μ та σ. Функція щільності нормального розподілу $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}$. Вивести формулу для функції нормального розподілу $F(x)$, математичного сподівання $M(x)$, дисперсії $D(x)$, ймовірності події $\alpha \leq X \leq \beta$, імовірності того, що значення випадкової величини відхилиться від математичного сподівання μ на величину, що не перевищує δ ($P(|x - \mu| \leq \delta)$).

14. Автомат виготовляє кульки. Кулька вважається стандартною, якщо відхилення X діаметра кульки від проектного розміру за абсолютною величиною менше, ніж 0,7 мм. Уважається, що НВВ X має нормальний розподіл із СКВ $\sigma = 0{,}4$ мм. Знайти, скільки в середньому буде стандартних кульок серед 100 виготовлених.

15. Випадкові похибки вимірювань мають нормальний розподіл із СКВ $\sigma = 20$ мм і математичним сподіванням $\mu = 0$. Знайти ймовірність того, що з трьох незалежних вимірювань похибка хоча б одного не перевищить за абсолютною величиною 4 мм.

16. Автомат штампує деталі. Контролюється довжина деталі X, що має нормальний розподіл з математичним сподіванням $\mu = 50$ мм (проектна

довжина). Фактична довжина виготовлених деталей не менша, ніж 32 та не більша, ніж 68 мм. Знайти ймовірність того, що довжина наугад взятої деталі: а) більше 55 мм; б) менша, ніж 40 мм.

17. Відомо, що середня кількість пакетів, яку генерує один з терміналів локальної мережі підприємства протягом години, має нормальний розподіл з параметрами $(10; 2)$. Знайти:

– імовірність того, що за годину кількість пакетів відхилиться від середнього менше, ніж на $\delta = 0.01$;

– імовірність того, що за годину кількість пакетів буде від 9 до 11;

– імовірність того, що за годину кількість пакетів буде < 9 або > 11.

Контрольні питання

1. Навести декілька прикладів дискретної випадкової величини.
2. Навести декілька прикладів неперервної випадкової величини.
3. Чи для всіх розподілів існують математичне сподівання і дисперсія?
4. Як виправдати використання математичного сподівання як числової характеристики для розподілу, який не має скінченного математичного сподівання?
5. Яка форма закону розподілу є універсальною і може бути застосовна як для ДВВ, так і для НВВ?
6. Які альтернативні числові характеристики можна використовувати для опису розподілу, якщо математичне сподівання не відображає його повністю?
7. У чому полягає ймовірнісний та статистичний сенс математичного сподівання?
8. Чому важливо враховувати асиметрію та ексцес під час аналізу розподілу величин?
9. Чому, якщо для певної ВВ не існує математичного сподівання, то не існує дисперсія, асиметрія і ексцес? Відповідь обґрунтуйте.
10. Чому на практиці часто можна апріорі вважати розподіл ВВ нормальним?

Література: [2, стор. 37–55, 75–94, 3, стор. 41–51, 57–68, 5, 7, 11, стор. 36-44, 58–71]

Практична робота № 6

Тема. Закони розподілу функцій випадкових величин. Композиція законів розподілу. Розподіл екстремальних значень.

Мета: набути практичних навичок розв'язання задач з обчислення функцій від випадкових величин, їх законів розподілу та числових характеристик.

Короткі теоретичні відомості

Перетворення випадкових величин

Перетворення однієї випадкової величини

Розглядатимемо тільки перетворення випадкових величин з абсолютно неперервними розподілами.

Нехай випадкова величина ξ має функцію розподілу $F_\xi(x)$ і щільність $f_\xi(x)$. Побудуємо за допомогою функції $g: R \to R$ випадкову величину $\eta = g(\xi)$. Необхідно знайти її функцію розподілу і, якщо існує, щільність розподілу.

Зауваження. Щільність розподілу випадкової величини $\eta = g(\xi)$ існує не за будь-яких функцій g. Так, якщо функція кусково-стала, то випадкова величина η має дискретний розподіл і щільності її розподілу не існує.

Теорема. Нехай ξ має функцію розподілу $F_\xi(x)$ і щільність розподілу $f_\xi(x)$, і константа a відмінна від нуля. Тоді випадкова величина $\eta = a\xi + b$ має щільність розподілу$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right)$.

Теорема. Нехай ξ має функцію розподілу $F_\xi(x)$ і щільність розподілу $f_\xi(x)$, і функція $g: R \to R$ монотонна. Тоді випадкова величина $\eta = g(\xi)$ має щільність розподілу $f_\eta(x) = \frac{1}{|g'(g^{-1}(x))|} f_\xi\left(g^{-1}(x)\right)$.

Тут $g^{-1}(x)$ – функція розподілу, зворотна до g, і $\frac{1}{g'(g^{-1}(x))} = \left(g^{-1}(x)\right)'$ - похідна функції $g^{-1}(x)$.

Висновки з теореми

Висновок 1. Якщо $\xi \sim N(0; 1)$, то $\eta = \sigma\xi + a \sim N(a; \sigma^2)$.

Висновок 2. Якщо $\eta \sim N(a; \sigma^2)$, то $\xi = \frac{\eta - a}{\sigma} \sim N(0; 1)$.

Висновок 3. Якщо $\eta \sim E(\alpha)$, то $\eta = \alpha\xi \sim E(1)$.

Приклад 6.1. Випадкова величина X розподілена за нормальним законом з нульовим математичним сподіванням. Знайти закон розподілу функції $Y = X^3$.

Розв'язання. Згідно з визначенням нормального розподілу неперервної випадкової величини X та умовою прикладу, диференціальний закон розподілу X має вигляд: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$. Функція $Y = X^3$ диференційована і зростає для $\forall x \in (-\infty, \infty)$, $Y' = 3X^2$. Тоді $Y = X^3 \Rightarrow X = \sqrt[3]{Y}$, тобто $\psi(y) = \sqrt[3]{Y} = y^{\frac{1}{3}}$. Отже матимемо $g(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y^{\frac{1}{3}})^2}{2\sigma^2}} \left| \left(y^{\frac{1}{3}} \right)' \right| = \frac{e^{-\frac{y^{\frac{2}{3}}}{2\sigma^2}}}{\sigma\sqrt{2\pi} \cdot y^{\frac{2}{3}}}$.

Функції від двох випадкових величин

Нехай ξ_1, ξ_2 – випадкові величини з щільністю сумісного розподілу $f_{\xi_1,\xi_2}(x_1, x_2)$, і задана функція $g: R^2 \to R$. Необхідно знайти функцію (а якщо існує, то і щільність) розподілу випадкової величини $\eta = g(\xi_1, \xi_2)$.

Теорема. Нехай $x \in R$ і область $D_x \subseteq R^2$ складається з точок (x_1, x_2), таких, що $g(x_1, x_2) < x$. Тоді випадкова величина $\eta = g(\xi_1, \xi_2)$ має функцію розподілу

$$F_\eta(x) = p\big((\xi_1, \xi_2) \in D_x\big) = \iint_{D_x} f_{\xi_1,\xi_2}(x_1, x_2)\, dx_1 dx_2.$$

Зауваження. Уважається, що випадкові величини ξ_1 і ξ_2 незалежні, тобто $f_{\xi_1,\xi_2}(x_1, x_2) = f_{\xi_1}(x_1) \cdot f_{\xi_2}(x_2)$.

Висновок (Формула згортки). Якщо випадкові величини ξ_1 і ξ_2 незалежні і мають абсолютно неперервний розподіл із щільностями $f_{\xi_1}(x_1)$ і $f_{\xi_2}(x_2)$, то щільність розподілу суми $\xi_1 + \xi_2$ дорівнює «згортці» щільностей $f_{\xi_1}(x_1)$ і $f_{\xi_2}(x_2)$:

$$f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} f_{\xi_1}(u) \cdot f_{\xi_2}(t-u)\, du = \int_{-\infty}^{\infty} f_{\xi_1}(t-u) \cdot f_{\xi_2}(u)\, du.$$

Приклад 6.2. Дві незалежні випадкові величини ξ і η мають стандартний нормальний розподіл: $\xi \sim N(0,1)$, $\eta \sim N(0,1)$. Покажемо, що $\xi+\eta \sim N(0,2)$.

Розв'язання.

$f_{\xi+\eta}(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi} \cdot e^{-\frac{1}{2}u^2} e^{-\frac{1}{2}(x-u)^2}\,du = \int_{-\infty}^{\infty} \frac{1}{2\pi} \cdot e^{-\frac{1}{2}(2u^2+x^2-2xu)}\,du =$

$= \int_{-\infty}^{\infty} \frac{1}{2\pi} \cdot e^{-(u^2+\frac{1}{2}x^2-xu)}\,du.$

Так як $u^2 + \frac{x^2}{2} - xu = \left(u - \frac{x}{2}\right)^2 + \frac{x^2}{4}$, то

$$f_{\xi+\eta}(x) = e^{-\frac{1}{2}x^2} \int_{-\infty}^{\infty} \frac{1}{2\pi} \cdot e^{-\frac{1}{2}\left(u-\frac{x}{2}\right)^2}\,du = e^{-\frac{1}{4}x^2} \int_{-\infty}^{\infty} \frac{1}{2\pi} \cdot e^{-v^2}\,dv =$$

$= \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}x^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \cdot e^{-v^2}\,dv = \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}x^2}.$

Закони розподілу екстремальних значень неперервної випадкової величини

Нехай $x_1, x_2, ..., x_n$ – вибірка неперервної випадкової величини X з функцією розподілу $F(x)$. Необхідно встановити закони розподілу $X_{\max} = \max(x_1, x_2, ..., x_n)$ і $X_{\min} = \min(x_1, x_2, ..., x_n)$.

Закон розподілу $X_{\max}$

За визначенням $F(X_{\max}) = p(X_{\max} < x)$. Тоді

$F(X_{\max}) = p(X_{\max} < x) = p(\max(x_1, x_2, ..., x_n) < x) =$

$= p(\max(x_1, x_2, ..., x_n) < x) = p(x_1 < x, x_2 < x, ..., x_n < x) =$

$$= p(x_1 < x) p(x_2 < x) ... p(x_n < x) = \prod_{i=1}^{n} p(x_i < x) = \prod_{i=1}^{n} \int_{-\infty}^{x} f(x)dx =$$

$$= \left(\int_{-\infty}^{x} f(x)dx \right)^n. \text{ Тобто}$$

$$F(X_{\max}) = \left(\int_{-\infty}^{x} f(x)dx \right)^n.$$

Тоді

$$f(X_{\max}) = F'(X_{\max}) = \left[\left(\int\limits_{-\infty}^{x} f(x)dx\right)^{n}\right]' = n\left(\int\limits_{-\infty}^{x} f(x)dx\right)^{n-1}\left(\int\limits_{-\infty}^{x} f(x)dx\right)' =$$

$$= n\left(\int\limits_{-\infty}^{x} f(x)dx\right)^{n-1} f(x).$$ Тобто

$$f(X_{\max}) = n\left(\int\limits_{-\infty}^{x} f(x)dx\right)^{n-1} f(x).$$

Закон розподілу $X_{\min}$

Аналогічно отримуємо

$$F(X_{\min}) = p(X_{\min} < x) = 1 - p(X_{\min} > x) = 1 - p(\min(x_1, x_2, ..., x_n) > x) =$$

$$= 1 - p(x_1 > x, x_2 > x, ..., x_n > x) = 1 - p(x_1 > x)p(x_2 > x)...p(x_n > x) =$$

$$= 1 - \prod_{i=1}^{n} p(x_i > x) = 1 - \prod_{i=1}^{n}\left(1 - p(x_i < x)\right) = 1 - \prod_{i=1}^{n}\left(1 - \int\limits_{-\infty}^{x} f(x)dx\right) =$$

$$= 1 - \left(1 - \int\limits_{-\infty}^{x} f(x)dx\right)^{n}.$$ Тобто

$$F(X_{\max}) = 1 - \left(1 - \int\limits_{-\infty}^{x} f(x)dx\right)^{n}.$$

Тоді $f(X_{\max}) = F'(X_{\max}) = \left[1 - \left(1 - \int\limits_{-\infty}^{x} f(x)dx\right)^{n}\right]' =$

$$= -n\left(1 - \int\limits_{-\infty}^{x} f(x)dx\right)^{n-1}\left(1 - \int\limits_{-\infty}^{x} f(x)dx\right)' = n\left(1 - \int\limits_{-\infty}^{x} f(x)dx\right)^{n-1} f(x).$$ Тобто

$$f(X_{\max}) = n\left(1 - \int_{-\infty}^{x} f(x)dx\right)^{n-1} f(x).$$

Задачі для самостійного розв'язання

Виконати індивідуальне завдання. Завдання полягає у розв'язанні п'яти задач, які потрібно вибрати зі списку, наведеного нижче. Правило вибору номерів таке: $n, n+1, n+2, n+3, n+4$, де n – номер студента у списку групи. У разі, якщо було досягнуто кінця списку задач, потрібно циклічно повернутися до його початку.

Завдання

1. Унаслідок експериментальних досліджень установлено, що похибка, зумовлена шумом аналогових елементів аналого-цифрового перетворювача (АЦП) комп'ютеризованої системи контролю (КСК), – випадкова величина X, що має нормальний закон розподілу з параметрами $\mu = 0, \sigma = 1$, а похибка квантування – випадкова величина Y, що має рівномірний розподіл з параметрами $a = 0, b = 1$. Знайти закон розподілу сумарної похибки $Z = X + Y$.

2. Випадкові величини X і Y незалежні та обидві мають рівномірний закон розподілу з параметрами $a = 0, b = 2$. Знайти функції розподілу та щільності розподілу випадкової величини $Z = X + Y$.

3. Час між запитами до сервера комп'ютерної мережі є випадковою величиною X, що має експоненціальний закон розподілу з параметром $\lambda = 10$. З метою дослідження степені використання сервера необхідно встановити закон розподілу максимумів випадкової величини X, тобто деякої випадкової величини $Z = max(X)$.

4. Випадкова величина X має рівномірний розподіл з параметрами $a = 0, b = \pi$. Знайти функції розподілу та щільності розподілу випадкової величини $Z = sin(X)$, обчислити математичне сподівання $M(Z)$ та дисперсію $D(Z)$.

5. Випадкова величина X має рівномірний розподіл з параметрами $a = 0, b = \pi$. Знайти функції розподілу та щільності розподілу випадкової величини $Z = cos(X)$, обчислити математичне сподівання $M(Z)$ та дисперсію $D(Z)$.

6. Час між запитами до сервера комп'ютерної мережі є випадковою величиною X, що має експоненціальний закон розподілу з параметром $\lambda = 5$. З метою дослідження степені використання сервера необхідно встановити закон розподілу мінімумів випадкової величини X, тобто деякої випадкової величини $Z = min(X)$.

7. Знайти закон розподілу $Z = X + Y$. $X \sim N(a;\sigma^2)$, $Y \sim E(\lambda)$.

8. Знайти закон розподілу $Z = X + Y$. $X \sim N(a_1;\sigma_1^2)$, $Y \sim N(a_2;\sigma_2^2)$.

9. Знайти закон розподілу $Z = X + Y$. $X \sim N(a;\sigma^2)$, $Y \sim U(a;b)$.

10. Знайти закон розподілу $Z = X + Y$. $X \sim U(a;b)$, $Y \sim U(a;b)$.

11. Знайти закон розподілу $Z = X + Y$. $X \sim U(-a;a)$, $Y \sim U(-\frac{a}{2};\frac{a}{2})$.

12. Знайти закон розподілу $Z = X + Y$. $X \sim E(\lambda_1)$, $Y \sim E(\lambda_2)$.

13. $X \sim U(0;2\pi)$. $Z = \sin X$. Знайти закон розподілу ВВ $Z = max(X)$.

9. $X \sim U(0;2\pi)$. $Z = \cos X$. Знайти закон розподілу ВВ $Z = max(X)$.

10. $X \sim U(0;\pi)$. $Z = \sin X$. Знайти закон розподілу ВВ $Z = max(X)$.

11. $X \sim U(0;\pi)$. $Z = \cos X$. Знайти закон розподілу ВВ $Z = min(X)$.

12. $X \sim U(a;b)$. Знайти закон розподілу ВВ $Z = \max(X)$.

13. $X \sim U(a;b)$. Знайти закон розподілу ВВ $Z = \min(X)$.

Контрольні питання

1. Як знання закону розподілу значень пікових навантажень у комп'ютерній мережі підприємства може допомогти у моделюванні та аналізі пікових навантажень?

2. Як знайти математичне сподівання функції одного випадкового аргумента?

3. Як знайти дисперсію функції одного випадкового аргумента?

4. Чому на етапі обчислення закону розподілу функції від випадкової величини потрібно виконати аналіз монотонності функції?

5. Наведіть приклади задач, де виникає потреба в обчисленні закону розподілу суми випадкових величин.

Література: [2, стор. 6–13, 11, стор. 11-14]

Практична робота № 7

Тема. Найпростіший потік подій. Елементи теорії СМО. Ланцюги Маркова

Мета: набути практичних навичок розв'язання задач щодо випадкових процесів, СМО та ланцюгів Маркова.

Короткі теоретичні відомості

Найпростіший потік подій

Визначення. Потоком подій називають послідовність подій, які надходять у випадкові моменти часу.

Головних властивості потоків такі.

1. *Властивість стаціонарності* (імовірність появи k подій за час t залежить тільки від k та t).

2. *Відсутність наслідків* (імовірність появи k подій протягом інтервалу часу $(T, T+t)$ не залежить від того, скільки подій і які існували до моменту T).

3. *Ординарність* (імовірність появи більше однієї події за малий проміжок часу Δt є нескінченно мала більш вищого порядку, ніж Δt, тобто, $P_{k>1}(\Delta t) = o(\Delta t)$, де $\lim\limits_{\Delta t \to 0} \dfrac{o(\Delta t)}{\Delta t} = 0$).

Визначення. Найпростішим (пуассонівським) називають потік подій, який має властивості стаціонарності, відсутності наслідків та ординарності.

Визначення. Інтенсивністю потоку λ називають середню кількість подій, які з'являються за одиницю часу.

Випадкові процеси

Процес Пуассона

Якщо події, що відбулися у випадкові моменти часу, відповідають вказаним вище властивостям 1–3, то кількість подій, що існують за фіксований проміжок часу t, розподілено за законом Пуассона з параметром λt. Інакше кажучи, якщо $X(t)$ – *кількість подій найпростішого потоку, що відбулися за час* t, то за зафіксованого t функція $X(t)$ є випадковою величиною, розподіленою за законом Пуассона з параметром λt.

Визначення. Функцію $X(t)$ називають *випадковим процесом Пуассона,* якщо при зафіксованому t функція $X(t)$ є випадкова величина, розподілена за законом Пуассона з параметром λt.

Процес Пуассона широко використовують для розв'язання багатьох задач практики і особливо в *теорії масового обслуговування*.

Зауваження. Інтервал часу між появами двох послідовних подій найпростішого потоку (випадкова величина T) розподілена за експоненціальним законом**,** тобто її функція розподілу має вигляд $F(t) = 1 - e^{-\lambda t}$.

Вінеровський процес (процес броунівського руху)

Визначення. Випадковий процес $X(t)$ називають *нормальним (гауссівским),* якщо сумісний розподіл $X(t_1), X(t_2), ..., X(t_k)$ є нормальним для $\forall k$ та для $\forall t_i, i = \overline{0,k}$.

Нормальний процес повністю визначається його математичним сподіванням і кореляційною функцією.

Визначення. Випадковий процес $X(t)$ називають процесом з *незалежними приростами*, якщо його прирости на інтервалах, що не перетинаються, взаємно незалежні, тобто випадкові величини $X(t_2)-X(t_1), X(t_3)-X(t_2),..., X(t_k)-X(t_{k-1})$ для $t_1 < t_2 < ... < t_k$ взаємно незалежні.

Визначення. Якщо $X(t)-X(s)$ залежить тільки від різниці $t-s$, то процес називають процесом *зі* стаціонарними *приростами.*

Визначення. Вінеровським процесом (процесом броунівського руху) називають нормальний випадковий процес $X(t)$ з незалежними стаціонарними приростами, для якого $X[0]=0$, $M[X(t)]=0$, $M[X(t)^2]=\sigma^2 t$ для $\forall t > 0$.

Марківський процес

Уведемо низку необхідних термінів. Нехай в кожний момент часу деяка система може знаходитись у одному зі станів $S_0, S_1,...$. (кількість станів скінченна, або зчислена). Якщо система переходить з одного, наприклад, S_i, у другий, наприклад, у S_j, то говорять, що у системі має місце випадковий процес.

Визначення. Якщо ймовірність переходу зі стану S_i у стан S_j залежить тільки від стану S_i і не залежить від того, коли і як система прийшла у цей стан, то випадковий процес називають *марківським*.

Визначення. Якщо для кожного моменту часу t_0 перебіг випадкового процесу $X(t)$ у майбутньому (при $t > t_0$) визначається поточним $X(t_0)$ і не залежить від минулого (значень $X(t)$ при $t < t_0$), то $X(t)$ – *марківський випадковий процес*.

Прикладом марківського процесу є процес обслуговування найпростішого потоку запитів *системою масового обслуговування* (*СМО*) з очікуванням (у такій системі запит «становиться у чергу», якщо усі канали зайняті) та експоненціальним часом обслуговування.

Граф станів СМО «n-клієнтів –Web-сервер»

Web-сервер комп'ютерної мережі на рис. 7.1 розглядається як прилад обслуговування стосовно до n - клієнтів: Клієнт 1, Клієнт 2., …, Клієнт n.

Кожний клієнт у середньому витрачає час $T_{кл}$ на прийняття, обробку та передавання одного запиту. Інтенсивність потоку завдань від одного клієнта дорівнює $\lambda = \dfrac{1}{T_{кл}}$. Усі клієнти можуть працювати паралельно, незалежно один від одного**,** але Web-сервер на базі однієї ЕОМ може працювати у режимі розподілу часу і у середньому витрачає час $T_{\text{ЕОМ}}$ на прийняття, обробку та передавання (фактично відсутній кеш, тобто *черги немає*). Тоді інтенсивність потоку інформації від ЕОМ до операторів дорівнює $\mu = \dfrac{1}{T_{EOM}}$. Уважатимемо, що завдання знаходиться у *клієнтській фазі*, якщо його обробляє оператор, і у *системній фазі*, якщо його обробляє Web-сервер. Стан СМО $S_j\,(j = \overline{1,n})$ визначається кількістю запитів клієнтів у системній фазі, що обробляє Web-сервер. Тобто ми маємо $(n+1)$ станів $S_0, S_1, ..., S_n$ з імовірністю існування $P_0, P_1, ..., P_n$. Ці стани означають:

S_0 – у системній фазі немає жодного завдання, усі завдання знаходяться у клієнтів (у клієнтській фазі);

S_1 – одне завдання знаходиться у системній фазі;

S_n – n завдань знаходиться у системній фазі.

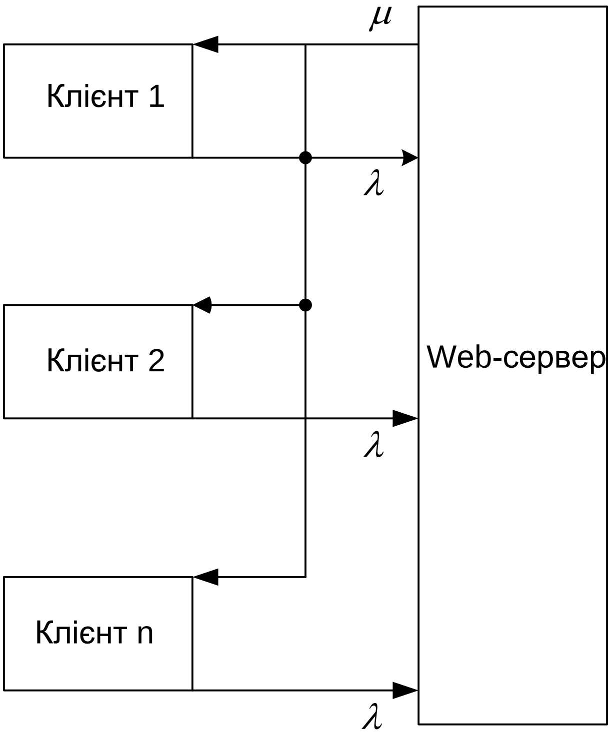

Рисунок 7.1 – СМО «n-клієнтів –Web-сервер» (система M/M/1)

Ця система є багатоканальною у разі введення інформації клієнтами і одноканальною під час обробки інформації.

Відповідний граф станів СМО має вигляд, наведений на рис. 3.2. Показаний граф станів має такі особливості.

1. Інтенсивність потоків переходу зі стану S_{j-1} до стану S_j (або з клієнтської фази у системну) визначається добутком $(n-j)\cdot\lambda$, тобто залежить від кількості клієнтів у клієнтській фазі $(n-j)$.

2. Інтенсивність потоків переходу зі стану S_j до стану S_{j-1} (або із системної фази у клієнтську) вважаємо постійною. Тобто вважаємо, що Web-сервер працює в режимі розподілу часу, тому інтенсивність зворотного потоку інформації однакова для всіх станів і його інтенсивність дорівнює $\mu = const$.

3. Для кожного зі станів S_j указуємо ймовірність його існування P_j.

Слід зазначити, що якщо б кожний клієнт мав свій окремий, незалежний канал обслуговування, то інтенсивність зворотного потоку інформації дорівнювала б добутку $j\mu$.

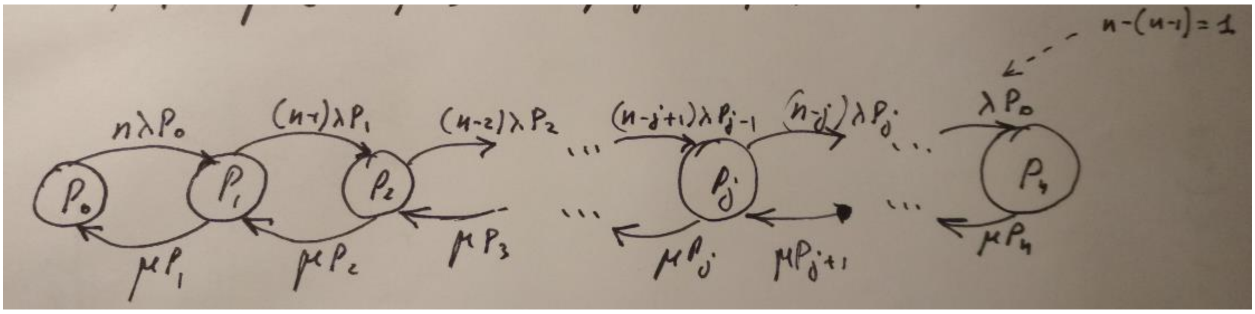

Рисунок 7.2 – Граф станів $S_0 - S_n$ СМО «n-клієнтів –Web-сервер»

Диференціальні рівняння і розрахунок імовірностей станів СМО

Систему диференціальних рівнянь складають за графом станів відносно невідомих ймовірностей P_j існування станів S_j. Для цього використовують такі правила.

1. Число рівнянь дорівнює $(n+1)$ – кількості станів S_j у графі.

2. У лівій частині рівняння стоїть похідна ймовірності за часом, а у правій частині – члени, пов’язані з дугами, які виходять та входять у вершину цього стану S_j.

Кожний член правої частини беруть зі знаком «+», якщо *входить* у вершину S_j, і зі знаком «–», якщо виходить з неї. Кожний член правої частини дорівнює добутку інтенсивності потоку інформації на ймовірність того стану, звідки *виходить дуга*.

Одне з отриманих рівнянь є зайвим і замінюється рівнянням суми ймовірностей повної групи для взаємно несумісних подій.

Із графу станів отримуємо рівняння:

$$\frac{dP_0}{dt} = -n\lambda P_0 + \mu P_1;$$

$$\frac{dP_1}{dt} = -(n-1)\lambda P_1 + \mu P_2 + n\lambda P_0 - \mu P_1;$$

$$\frac{dP_2}{dt} = -(n-2)\lambda P_2 + \mu P_3 + (n-1)\lambda P_1 - \mu P_2;$$

…………………………………………….

$$\frac{dP_n}{dt} = \lambda P_{n-1} - \mu P_n.$$

Останнє рівняння замінюємо на суму ймовірностей повної групи:

$$\sum_{j=0}^{n} P_j = 1.$$

Далі отримуємо іншу систему рівнянь:

$$\frac{dP_0}{dt} = -n\lambda P_0 + \mu P_1 = 0;$$

$$\frac{dP_0}{dt} + \frac{dP_1}{dt} = -(n-1)\lambda P_1 + \mu P_2 = 0;$$

$$\frac{dP_0}{dt} + \frac{dP_1}{dt} + \frac{dP_2}{dt} = -(n-2)\lambda P_2 + \mu P_3 = 0;$$

…………………………………………….

$$\frac{dP_0}{dt} + \frac{dP_1}{dt} + ... + \frac{dP_j}{dt} = -(n-j)\lambda P_j + \mu P_{j+1} = 0;$$

$$\sum_{j=0}^{n} P_j = 1.$$

У статичному режимі всі похідні дорівнюють нулю, і тому отримуємо:

$$n\lambda P_0 = \mu P_1;\ (n-1)\lambda P_1 = \mu P_2;\ \ldots;$$

$$(n-j)\lambda P_j = \mu P_{j+1};\ \sum_{j=0}^{n} P_j = 1, \text{ де } j = \overline{0,n}.$$

Розв’язання цієї системи рівнянь для статики має вигляд:

$$P_1 = \frac{n\lambda P_0}{\mu};\ P_2 = \frac{(n-1)\lambda P_1}{\mu} = \frac{(n-1)\lambda^2 P_0}{\mu^2};\ldots;$$

$$P_j = (n-j+1)...(n-2)(n-1)n\left(\frac{\lambda}{\mu}\right)^j = \frac{n!}{(n-j)!}\left(\frac{\lambda}{\mu}\right)^j P_0.$$

Звідси отримуємо ймовірність простою СМО (коли кількість вимог $i = 0$):

$$1 = \sum_{j=0}^{n} P_i = P_0 \sum_{j=0}^{n} \frac{n!}{(n-j)!}\left(\frac{\lambda}{\mu}\right)^j;$$

$\Rightarrow$

$$P_0 = \frac{1}{\sum_{j=0}^{n} \frac{n!}{(n-j)!}\left(\frac{\lambda}{\mu}\right)^j}.$$

Web-сервер зайнятий обробкою запитів (вимог) клієнтів з імовірністю

$P_{зан} = 1\text{-}P_0$,

де P_0 – ймовірність відсутності запитів.

Якщо Web-сервер зайнятий, то він обслуговує μ програм в одиницю часу. Тому пропускна спроможність Web-сервера інформаційної системи дорівнює:

$$A = (1 - P_0)\mu.$$

Середній потік запитів до Web-сервера:

$$A = \lambda(n - w),$$

де $(n-w)$ – середня кількість клієнтів, що знаходяться у клієнтській фазі; w – середня кількість клієнтів у системній фазі.

Цей потік повинен дорівнювати пропускній спроможності Web-сервера, і тому:

$$(1 - P_0)\mu = \lambda(n - w),$$

звідки середня кількість клієнтів у системній фазі:

$$w = n - (1 - P_0)\frac{\mu}{\lambda}.$$

На кожній з w запитів Web-сервер витрачає $T_{обс}$. Тому середній час відгуку Web-сервера $T_{відг} = wT_{обс}$.

Розраховані дані дозволяють скласти вимоги до параметрів Web-сервера.

Ланцюг Маркова

Визначення. Ланцюгом Маркова називають послідовність випробувань, у кожному з яких відбувається тільки одна з k несумісних подій $A_1, A_2, \dots, A_k$ повної групи, причому умовна ймовірність $p_{js}(s)$ того, що в s-му випробуванні відбудеться подія $A_j, j = 1, 2, \dots, k$, за умови, що в $(s-1)$-му випробуванні відбулася подія $A_i, i = 1, 2, \dots, k$, не залежить від результатів попередніх випробувань.

Ланцюгом Маркова з *дискретним* часом називають ланцюг, зміна стану якого відбувається в певні *фіксовані* моменти часу.

Однорідний ланцюг Маркова. Перехідні ймовірності. Матриця переходу

Визначення. Однорідним називають ланцюг Маркова, якщо умовна ймовірність $p_{ij}(s)$ (переходу з стану i у стан j) не залежить від номера випробування. Тому писатимемо просто p_{ij}.

Визначення. Перехідною ймовірністю p_{ij} називають умовну ймовірність того, що зі стану i (у якому система опинилася в результаті деякого випробування, байдуже якого номера) в результаті наступного випробування система перейде у стан j.

Визначення. Матрицею переходу системи називають матрицю, яка містить всі перехідні ймовірності цієї системи: $P_1 = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{pmatrix}$,

$\sum_{j=1}^{k} p_{ij} = 1 \ (i = 1, 2, \dots, k)$.

Можна показати, що у разі, коли матриця переходів відома, імовірність переходу системи зі стану i у стан j за n кроків може бути обчислена так: $P_n = P_1^n$.

Приклад 7.1. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{pmatrix}$. Знайти матрицю переходу P_2.

Розв'язання. Скористаємося формулою $P_n = P_1^2$. $P_1 = \begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.34 & 0.66 \\ 0.33 & 0.67 \end{pmatrix}$.

Задачі для самостійного розв'язання

Виконати індивідуальне завдання. Завдання полягає у розв'язанні двох задач, які потрібно вибрати зі списку, наведеного нижче. Правило вибору номерів таке: $n, n+1$, де n – номер студента у списку групи. У разі, якщо було досягнуто кінця списку задач, потрібно циклічно повернутися до його початку.

Завдання

1. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 2$, $\lambda = 1$, $\mu = 2$. $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

2. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{pmatrix}$. Знайти матрицю переходу P_2.

3. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 2$, $\lambda = 2$, $\mu = 1$ і знайти ймовірності станів. $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

4. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix}$. Знайти матрицю переходу P_2.

5. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 4$, $\lambda = 1$, $\mu = 1$. $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

6. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$. Знайти матрицю переходу P_2.

7. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 4$, $\lambda = 1$, $\mu = 2$. $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

8. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix}$. Знайти матрицю переходу P_2.

9. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 4$, $\lambda = 2$ $\mu = 1$. Знайти $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

10. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{pmatrix}$. Знайти матрицю переходу P_2.

11. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 5$, $\lambda = 2$ $\mu = 1$. Знайти $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

12. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$. Знайти матрицю переходу P_2.

13. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 5$, $\lambda = 1$ $\mu = 2$. Знайти $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

14. 7. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$. Знайти матрицю переходу P_3.

15. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 3$, $\lambda = 1$ $\mu = 2$. Знайти $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

16. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix}$. Знайти матрицю переходу P_3.

17. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 2$, $\lambda = 1$ $\mu = 2$. Знайти $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

18. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{pmatrix}$. Знайти матрицю переходу P_3.

19. Побудувати граф станів СМО «n-клієнтів –Web-сервер» (система M/M/1) і систему рівнянь Колмогорова для $n = 3$, $\lambda = 2$ $\mu = 2$. Знайти $P_{зан}, P_0, A, w, T_{обс}, T_{відг}$.

20. Задано матрицю переходу $P_1 = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$. Знайти матрицю переходу P_3.

Контрольні питання

1. Що таке СМО і які головні елементи є у її структурі?

2. Які властивості має найпростіший потік подій, і які його характеристики можна виміряти?

3. Які основні характеристики СМО визначають її продуктивність?

4. Які чинники впливають на інтенсивність потоку подій у системі масового обслуговування?

5. Як визначається інтенсивність обслуговування в СМО?

6. Які властивості мають ланцюги Маркова, і як їх застосовують у теорії СМО?

7. Що таке стаціонарний режим роботи СМО і чому він важливий для аналізу?

8. Як визначається ймовірність утрати заявки в системі масового обслуговування?

9. Що таке ефективність обслуговування в СМО і як її вимірюють?

10. Як визначається коефіцієнт завантаження системи масового обслуговування, і чому він важливий для оцінки її продуктивності?

Література: [14, 15]

Практична робота № 8

Тема. Основи вибіркового методу. Точкові статистичні оцінки. Точні вибіркові розподілення. Інтервальні статистичні оцінки.

Мета: набути практичних навичок розв’язання типових задач з основ вибіркового методу, точкового та інтервального оцінювання числових характеристик випадкової величини.

Короткі теоретичні відомості

Вибірка. Варіаційний ряд. Статистичний розподіл і гістограма. Міри центральної тенденції та розсіювання

Визначення. Вибіркою об’єму n називають випадковий вектор (одновимірний масив) $X = (x_1, x_2, ..., x_n)$, де $x_1, x_2, ..., x_n$ – незалежні і однаково розподілені випадкові величини з функцією розподілення $F(x)$.

Приклад 8.1. Маса тіла (*кг*) п’яти навмання вибраних студенток першого курсу: $X = (50, 65, 55, 55, 60)$.

Приклад 8.2. Пульс спортсмена (*кількість ударів за хв.*) на початку щоденного тренування протягом п’яти тижнів фіксувався по понеділках. Результати подано вектором $X = (72, 72, 64, 68, 72)$.

Головна задача *вибіркового методу* полягає в тому, щоб визначити статистичні характеристики випадкової величини, що вивчається, виключно на підставі інформації, яку містить вибірка внаслідок «згортання» інформації до невеликого набору числових характеристик і графіків. Інша інформація відсутня. Уважається, що на підставі вибірки буде зроблено висновок щодо поведінки всієї генеральної сукупності.

Визначення. Відсортовану за зростанням вибірку $\alpha_1, \alpha_2, ..., \alpha_k$, $k \le n$, називають *варіаційним рядом.*

Визначення. Нехай $x_1, x_2, ..., x_n$ – вибірка з X об’єму n, $\alpha_1, \alpha_2, ..., \alpha_k$ – варіаційний ряд. Тоді число $\omega_i = \frac{m_i}{n} (i = 1,2,...,k)$, де m_i – кількість повторень варіанти α_i у вибірці об’єму n, називають *частотою* цієї *варіанти.*

Приклад 8.3. $X = (50,65,55,55,60)$.

Варіаційний ряд: $\alpha = (50,55,60,65)$.

Визначення. Нехай X – дискретна випадкова величина. Тоді таблицю

X	α_1	α_2	...	α_k
ω	ω_1	ω_2	...	ω_k

називають *таблицею відносних частот* або *емпіричним законом розподілення*. Графік – *полігон*.

Приклад 8.4. $X = (50,65,55,55,60)$.

X	50	55	60	65
ω	$1/5$	$2/5$	$1/5$	$1/5$

Зауваження. Якщо випадкова величина неперервна, то складаємо *інтервальний варіаційний ряд*:

1. Знаходимо $X_{\min}$ і $X_{\max}$ та увесь проміжок ділимо на k частин, яке, наприклад, визначається за формулою Стерджесса: $k = 1 + 1{,}332 \, ln \, n$.

2. Будуємо інтервальний варіаційний ряд: $[C_1; C_2[, [C_2; C_3[, ..., [C_k; C_{k+1}\ ]$

Визначення. Таблицю:

X	$[C_1; C_2[$	$[C_2; C_3[$	...	$[C_k; C_{k+1}\ ]$
ω	ω_1	ω_2	...	ω_k

називають *інтервальною таблицею частот*. Графік – *гістограма* (по осі $y - \frac{\omega}{h}, h = \frac{X_{\max} - X_{\min}}{k}$).

Приклад 8.5. $X = (50,65,55,55,60)$.

X	[50,55)	[55,60)	[60,65]
ω	$1/5$	$2/5$	$2/5$

Визначення. Емпіричною функцією розподілення, побудованою за вибіркою $X=(x_1,x_2,...,x_n)$ об'єму n, називають залежність відносної частоти ω події $X>x$ від x, що дорівнює $F_n^*(x)=\dfrac{кільк.x_i\in(-\infty;x)}{n}$.

Приклад 8.6.

X	50	55	60	65
ω	$1/5$	$2/5$	$1/5$	$1/5$

.

$$F_n^*(x)=\begin{cases}0, & x\le 50\\ 1/5 & 50<x\le 55\\ 1/5+2/5=3/5 & 55<x\le 60\\ 1/5+2/5+1/5=4/5 & 60<x\le 65\\ 1/5+2/5+1/5+1/5=5/5=1 & x>65\end{cases}$$

Міри центральної тенденції, розсіювання, форми розподілу та центилі

Для оцінювання середнього значення випадкової величини використовують так звані *міри центральної тенденції*.

Визначення. Вибірковою медіаною (median) випадкової величини називають середину варіаційного ряду, яку визначають за формулою:

$$\tilde{Me}=\begin{cases}\dfrac{1}{2}\left(\alpha_{\frac{n}{2}}+\alpha_{\frac{n}{2}+1}\right), n-парне\\ \alpha_{\frac{n+1}{2}}, n-непарне.\end{cases}$$

Медіана є робастною (нечутливою до викидів) характеристикою міри центральної тенденції.

Приклад 8.7.

X	50	55	60	65
n_i	1	2	1	1
ω	$1/5$	$2/5$	$1/5$	$1/5$

,

$$\tilde{Me}=\frac{1}{2}\left(\alpha_{\frac{4}{2}}+\alpha_{\frac{4}{2}+1}\right)=\frac{\alpha_2+\alpha_3}{2}=\frac{55+60}{2}=57{,}5.$$

Визначення. Середнім арифметичним (average, mean) (вибірковим математичним сподіванням) називають величину

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{i=1}^{m} \alpha_i n_i = \sum_{i=1}^{m} x_i \omega_i = \alpha^T \omega .$$

Приклад 8.8. $\bar{x} = \frac{1}{n}\sum_{i=1}^{m_i} \alpha_i n_i = \frac{1}{5}\left(50\cdot 1 + 55\cdot 2 + 60\cdot 1 + 65\cdot 1\right) =$

$$\alpha^T \omega = (50{,}55{,}65{,}60)\begin{pmatrix} 1/5 \\ 2/5 \\ 1/5 \\ 1/5 \end{pmatrix} = 57$$

Середнє арифметичне є чутливим до викидів, за наявності яких відбувається суттєве зміщення в оцінці середнього в бік викиду.

Визначення. Мода (moda) – значення варіаційного ряду, яке зустрічається найчастіше, тобто відповідає максимальному значенню частоти чи відносної частоти. Визначається за формулою:

$$Mo = \alpha_i(\max[\omega_i]).$$

Приклад 8.9. Статистичний розподіл: $\begin{matrix} X & 50 & 55 & 60 & 65 \\ \omega & 1/5 & 2/5 & 1/5 & 1/5 \end{matrix}$

$Mo = \alpha_i(\max[\omega_i]) = \alpha_2(\max[\omega_2 = 2/5]) = 55.$

Існує ще низка різного роду мір центральної тенденції. Що застосовуються у окремих випадках, але рідше, ніж вищезазначені.

Інша група числових характеристик призначена для оцінювання ступеня розсіювання значень випадкової величини – *міри розсіювання*. Найпростіша з них – *розмах (range).*

Визначення. Детерміновану величину $R = X_{\max} - X_{\min}$ називається *розмахом вибірки.*

Приклад 8.10. $X = (50{,}65{,}55{,}55{,}60)$. $R = X_{\max} - X_{\min} = 65 - 50 = 15$

Визначення. Виправлена вибіркова дисперсія (variance) – це середнє арифметичне квадрата відхилення від середнього арифметичного:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left(X^T X - n\bar{x}^2\right), \text{ або}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{m}(\alpha_i - \bar{x})^2 n_i = \frac{n}{n-1}\left((\alpha \circ \alpha)^T \omega - \bar{x}^2\right).$$

Приклад 8.11. $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{5-1}\sum_{i=1}^{n}(x_i - 57)^2 =$

$$= \frac{1}{4}\left[(50-57)^2 + (65-57)^2 + (55-57)^2 + (55-57)^2 + (60-57)^2\right] = 32{,}5$$

Визначення. Вибіркове середнє *арифметичне відхилення* і *виправлене вибіркове середнє арифметичне відхилення (standart deviation)* відповідно:

$$s = \sqrt{s^2}\,.$$

Приклад 8.12. $s = \sqrt{s^2} = \sqrt{32{,}5} = 5{,}701$

Визначення. Вибіркова середня *абсолютна похибка (Mean Absolute Error)*:

$$MAE = \frac{1}{n}\sum|x_i - \bar{x}|.$$

Приклад 8.13. $MAE = \frac{1}{n}\sum|x_i - \bar{x}| = \frac{1}{5}\sum|x_i - 57| =$

$$= \frac{1}{5}\left[|50-57| + |65-57| + |55-57| + |55-57| + |60-57|\right] =$$

$$\frac{1}{5}\left[|-7| + |8| + |-2| + |-2| + |3|\right] = \frac{1}{5}\left[7+8+2+2+3\right] = 4{,}4$$

Існує багато і інших числових характеристик.

Також існують числові характеристики, що характеризують *міру форми* розподілу випадкової величини. Головні з них – це *асиметрія (skewness)* та *ексцес (kurtosis)*.

Визначення. Вибіркова асиметрія:

$$\widetilde{A}_s = \frac{n\sum_{i=1}^{n}(x_i - \overline{x})^3}{(n-1)(n-2)s^3}.$$

Приклад 8.14. $\widetilde{A}_s = \frac{n\sum_{i=1}^{n}(x_i - \overline{x})^3}{(n-1)(n-2)s^3} = \frac{5\sum_{i=1}^{5}(x_i - 57)^3}{(5-1)(5-2)\cdot 5{,}701^3} =$

$$= \frac{5\left[(50-57)^3 + (65-57)^3 + (55-57)^3 + (55-57)^3 + (60-57)^3\right]}{4\cdot 3\cdot 185{,}279} = 0{,}405.$$

Значення асиметрії в околі нуля ($\widetilde{A}_s \approx 0$), указує на симетричність розподілу. Якщо $\widetilde{A}_s > 0$ – розподіл має правий «хвіст» більший, ніж лівий, що вказує на наявність окремих суттєвих викидів вправо, або на наявність суттєвого домінування значень, які вище, ніж мода, над значеннями, що є нижчими, ніж мода. Якщо $\widetilde{A}_s < 0$ – навпаки.

Визначення. Стандартизована вибіркова асиметрія (standartized skewness):

$$z_1 = \frac{\widetilde{A}_s}{\sqrt{6/n}}.$$

Приклад 8.15. $z_1 = \frac{\widetilde{A}_s}{\sqrt{6/n}} = \frac{0{,}405}{\sqrt{6/5}} = 0{,}37.$

Якщо значення z_1 виходять за межі інтервалу [-2,2], то є підстави вважати, що розподіл суттєво відрізняється від нормального.

Визначення. Вибірковий ексцес (kurtosis):

$$\widetilde{E}_k = \frac{n(n+1)\sum_{i=1}^{n}(x_i - \overline{x})^4}{(n-1)(n-2)(n-3)s^4} - \frac{3(n-1)^2}{(n-2)(n-3)}.$$

Приклад 8.16. $\tilde{E}_k = \dfrac{n(n+1)\sum\limits_{i=1}^{n}(x_i-\bar{x})^4}{(n-1)(n-2)(n-3)s^4} - \dfrac{3(n-1)^2}{(n-2)(n-3)} =$

$$= \frac{5(5+1)\sum\limits_{i=1}^{n}(x_i-57)^4}{(5-1)(5-2)(5-3)5{,}701^4} - \frac{3(5-1)^2}{(5-2)(5-3)} = -0{,}178$$

Якщо $\tilde{E}_k \approx 0$ – розподіл має купелоподібну форму, наближену до нормального. Якщо $\tilde{E}_k > 0$ – розподіл має більш загострену вершину у центрі, ніж у нормального, і більші за довжиною «хвости». Якщо $\tilde{E}_k < 0$ – розподіл має більш плоску вершину і коротші «хвости», ніж у нормального. Ця характеристика є *релевантною виключно для симетричних розподілів*.

Визначення. Стандартизований вибірковий ексцес (standartized kurtosis):

$$z_2 = \frac{E_k}{\sqrt{24/n}}.$$

Приклад 8.17. $z_2 = \dfrac{E_k}{\sqrt{24/n}} = \dfrac{-0{,}178}{\sqrt{24/5}} = -0{,}081.$

Якщо значення z_2 виходять за межі інтервалу [-2,2], то є підстави вважати, що розподіл суттєво відрізняється від нормального.

Визначення. Центиль рівня p-% – це значення варіаційного ряду α_i, лівіше якого знаходиться p-% усіх вибіркових значень.

Приклад 8.18. Медіана (показано вище) є 50-ти відсотковим центилєм, оскільки ділить варіаційний ряд рівно навпіл, тобто ліворуч і праворуч знаходиться рівно половина вибіркових значень.

Точкові та інтервальні оцінки параметрів нормального розподілення. П'ятиквантильний графік

Як точкова оцінка параметра a використовується $\bar{x} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i$.

Для інтервальної оцінки:

$\bar{x}-\dfrac{t(\gamma,n)s}{\sqrt{n}}<a<\bar{x}+\dfrac{t(\gamma,n)s}{\sqrt{n}}$, де $t(\gamma,n)$ – квантиль розподілу Стьюдента рівня γ, який знаходиться за спеціальними таблицями.

Як оцінка параметра σ^2 використовується виправлена вибіркова дисперсія s^2.

Приклад 8.19. $X=(50,65,55,55,60)$. $\bar{x}=57$, $s=\sqrt{s^2}=\sqrt{32,5}=5,701$, $t(0,95;5)=2,571$.

$$57-\frac{2,571\cdot 5,701}{\sqrt{5}}<a<57+\frac{2,571\cdot 5,701}{\sqrt{5}}$$

$$50,445<a<63,555.$$

Інтервальне оцінювання здійснюється за формулою:

$\dfrac{(n-1)s^2}{\chi^2_{n-1,\frac{1+\gamma}{2}}}<\sigma^2<\dfrac{(n-1)s^2}{\chi^2_{n-1,\frac{1-\gamma}{2}}}$, де $\chi^2_{n-1,\frac{1+\gamma}{2}}$ і $\chi^2_{n-1,\frac{1-\gamma}{2}}$ – квантилі рівня $\dfrac{1+\gamma}{2}$ та $\dfrac{1-\gamma}{2}$ розподілу χ^2 (знаходяться за таблицями).

Приклад 8.20. $X=(50,65,55,55,60)$, $s=\sqrt{s^2}=\sqrt{32,5}=5,701$,

$$\chi^2_{n-1,\frac{1+\gamma}{2}}=\chi^2_{4,\frac{1+0,95}{2}}=11,1,\quad \chi^2_{n-1,\frac{1-\gamma}{2}}=\chi^2_{4,\frac{1-0,95}{2}}=0,484$$

$$\frac{4\cdot 32,5}{11,1}<\sigma^2<\frac{4\cdot 32,5}{0,484},$$

$$11,712<\sigma^2<268,595$$

$$\sqrt{\frac{4\cdot 32,5}{11,1}}<\sigma<\sqrt{\frac{4\cdot 32,5}{0,484}},$$

$$3,422<\sigma<16,389.$$

Альтернативним варіантом графічного зображення цілісної інформації стосовно розподілу, поширеним у статистичних пакетах, є так званий *п'ятиквантильний графік*, або «ящик з вусами» (*Box-and-Wiskers Plot*) (рис. 4.1).

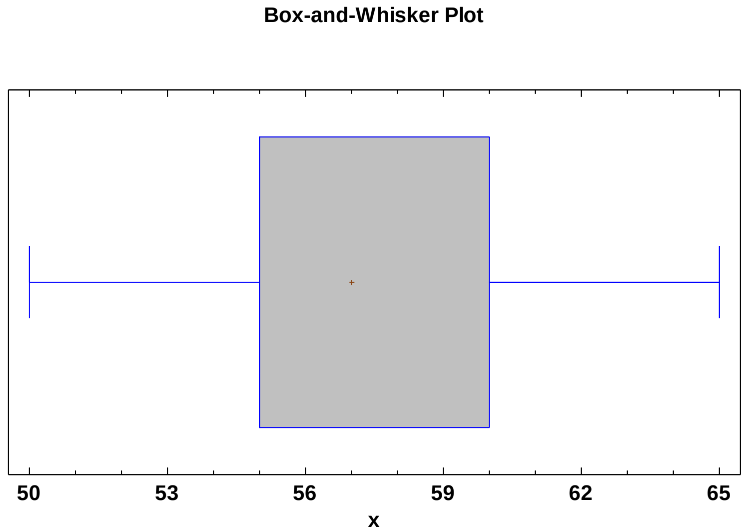

Рисунок 8.1 – П’ятиквантильний графік «ящик з вусами»

Сторони ящика відповідають нижньому і верхньому квартилям відповідно (25 %-му і 75 %-му центилям). Вертикальна лінія всередині ящика – це медіана (50-ти %-вий центиль). Зірочка – середнє арифметичне. «Вуса» відповідають найменшому і найбільшому значенню вибірки за винятком тих, що перевищують міжквартільну відстань за метрикою Тьюкі більше, ніж у півтора рази. Якщо значення викидів перевищують вищезазначену відстань більш, ніж у три рази, вони позначені зірочками.

Задачі для самостійного розв’язання

Виконати індивідуальне завдання, яке полягає в розрахунку числових і функціональних характеристик вибірки (вибірка вибирається згідно з варіантом) згідно з прикладами 8.1–8.20. Номер варіанта – номер студента у списку групи. У разі, якщо було досягнуто кінця списку задач, потрібно циклічно повернутися до його початку.

Завдання

1. [4 5 7 8 10]
2. [1 7 7 10 5]

3. [5 9 8 8 8]
4. [2 7 7 4 5]
5. [1 1 1 6 4]
6. [4 4 4 2 5]
7. [8 3 6 10 2]
8. [10 1 10 8 6]
9. [5 4 4 9 5]
10. [8 1 3 8 10]
11. [5 10 9 2 6]
12. [5 8 10 7 8]
13. [2 6 9 9 7]
14. [10 5 4 5 2]
15. [3 9 5 9 3]
16. [6 4 8 4 4]
17. [8 8 10 9 3]
18. [3 4 3 5 10]
19. [2 6 4 1 6]
20. [7 7 3 1 2]
21. [1 4 8 4 4]
22. [8 4 10 9 3]
23. [3 4 9 5 10]
24. [2 6 4 1 11]
25. [7 8 3 1 2]
26. [6 7 8 4 4]
27. [8 8 15 9 3]
28. [3 4 1 5 10]
29. [2 1 4 1 6]
30. [5 7 3 1 2]

Контрольні питання

1. Що таке вибірковий метод і як його використовують у статистиці?

2. Які є головні точкові статистичні оцінки, і як їх обчислюють?

3. Які чинники впливають на точність статистичних оцінок?

4. Як визначають вибіркову середню і вибіркову дисперсію?

5. Що таке точні вибіркові розподілення і як вони допомагають у роботі з вибірковими оцінками?

6. Які властивості мають інтервальні статистичні оцінки?

7. Як будується довірчий інтервал для параметра генеральної сукупності?

8. Як визначається довірчий інтервал для середнього значення генеральної сукупності?

9. Як використовувати інтервальні статистичні оцінки для прийняття рішень?

10. Які методи можна використовувати для визначення обсягу вибірки для отримання точних статистичних оцінок?

Література: [2, стор. 37–44, 9, 10, 11, стор. 31-36]

3 КРИТЕРІЇ ОЦІНЮВАННЯ ЗНАНЬ СТУДЕНТІВ

Практичні роботи студенти виконують у межах змістового модуля «Теорія ймовірностей та ймовірнісні процеси».

Усього студент повинен виконати вісім практичних завдань. За виконання всіх індивідуальних завдань у межах кожної практичної роботи студент може отримати максимально 40 балів. Максимальна кількість балів, яку можна отримати студенти за виконання однієї практичної роботи, складає 5 балів: 3 – «задовільно», 4 – «добре», 5 – «відмінно».

За кожне індивідуальне завдання в рамках практичної роботи виставляють:

– 5 балів, якщо студент виконав практичну роботу, написав звіт і захистив його, тобто правильно та вичерпно відповів на всі питання. Допускаються незначні неточності під час відповіді на одне з контрольних питань;

– 4 бали, якщо студент виконав практичну роботу, написав звіт і захистив його, але мав незначні помилки;

– 3 бали, якщо студент виконав практичну роботу, написав звіт і захистив його, але під час захисту не відповів на декілька питань і мав незначні помилки в оформленні звіту.

Розподіл балів за видами занять що отримують студенти, наведено у навчальній робочій програмі.

СПИСОК ЛІТЕРАТУРИ

1. Dembo A. (2019, April 23). Probability Theory: STAT310/MATH230. Department of Mathematics, Stanford University. Email: amir@math.stanford.edu. URL: https://web.stanford.edu/class/stats310a/lnotes.pdf

2. McMullen C. (2011). *Probability Theory: Course Notes.* Harvard University. Retrieved March 29, 2021. URL: https://people.math.harvard.edu/~ctm/papers/home/text/class/harvard/154/course/course.pdf

3. F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, L. E. A Meester. Modern Introduction to Probability and Statistics: Understanding Why and How Meester. Springer Science & Business Media, 2005 - 486 стор. URL: https://cis.temple.edu/~latecki/Courses/CIS2033-Spring13/Modern_intro_probability_statistics_Dekking05.pdf

4. Glass G. V., & Stanley J. C. (1970). Statistical Methods in Education and Psychology (Hardcover). Prentice Hall. (596 pages). Посилання на веб-сайт: https://www.amazon.com.au/Statistical-Methods-Education-Psychology-Glass/dp/0138449287

5. David F. Anderson, Timo Seppäläinen, Benedek Valkó. Introduction to Probability. Cambridge University Press, 2 лист. 2017 р.

6. Runyon R. P. (1977). *Nonparametric Statistics: A Contemporary Approach* (Addison-Wesley Second Language Professional Library Series) (Multilingual Edition). Addison Wesley Publishing Company. (218 pages) Посилання на веб-сайт: https://www.amazon.com/Nonparametric-statistics-contemporary-approach-Addison-Wesley/dp/0201065479

7. Mark Ward, Ellen Gundlach ,W. H. Freeman. Introduction to Probability. 12 черв. 2015 р. – 704 стор.

8. Draper N. R., & Smith H. (1998). Applied Regression Analysis (Wiley Series in Probability and Statistics) (Third Edition). Wiley-Interscience. (736 pages).

Посилання на веб-сайт: https://www.amazon.com/Applied-Regression-Analysis-Probability-Statistics/dp/0471170828

9. Stroock D. W. (2011). *Probability Theory: An Analytic View* (2nd ed.). Cambridge University Press. ISBN-10: 0521132509, ISBN-13: 978-0521132503.

10. MIT OpenCourseWare. (2014). *Theory of Probability* (Spring 2014). Instructor: Prof. Scott Sheffield. Department: Mathematics. Course URL: https://ocw.mit.edu/courses/18-175-theory-of-probability-spring-2014/

11. Найко Д. А. Шевчук О. Ф. Теорія ймовірностей та математична статистика: навч. посіб. Вінниця: ВНАУ, 2020. 382 с. http://repository.vsau.org/getfile.php/24513.pdf

12. Сидоренко В. М., Садовнича С. А., Долударєва Є. В. Оптимізація структури тестових завдань навчальних онлайн-курсів на основі ймовірнісної моделі / Інженерні та освітні технології. 2022. Т. 10. № 2. С. 27–36. doi: https://doi.org/10.30929/2307- 9770.2022.10.02.03

13. Сидоренко В. М., Кирилаха Н. Г. Дидактико-методичні аспекти викладання теорії ймовірностей та математичної статистики студентам IT напряму. Інженерні та освітні технології. 2023. Т. 11. № 3. С. 17–23. Doi: https://doi.org/10.32782/2307- 9770.2023.11.03.02

14. Sundarapandian, V. (2009). «7. Queueing Theory». Probability, Statistics and Queueing Theory. PHI Learning. ISBN 978-81-203-3844-9.

15. Kendall D. G. (1953). «Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain». The Annals of Mathematical Statistics. 24 (3): 338–354. doi:10.1214/aoms/1177728975. JSTOR 2236285.

Інформаційні ресурси

16. git. URL: https://git-scm.com/downloads (дата звернення: 14.07.2023).

17. git (укр.). URL: https://git-scm.com/book/uk/v2 (дата звернення: 14.07.2023).

18. Teacher's DevOps Course (Lecture 10). URL: https://www.youtube.com/watch?v=Md8RW6tKCNg (дата звернення: 14.07.2023).

19. Callout Blocks. Markdown Syntax. URL: https://quarto.org/docs/authoring/callouts.html (дата звернення: 14.07.2023).

20. BibTeX (Вікіпедія). URL: https://quarto.org/docs/authoring/callouts.html (дата звернення: 14.07.2023).

21. Literate prigramming (Вікіпедія). URL: https://en.wikipedia.org/wiki/Literate_programming (дата звернення: 14.07.2023).

22. Posit. (n.d.). Put data into production with Posit Connect. URL: https://posit.co/ (дата звернення: 14.07.2023)

23. The R Project for Statistical Computing. URL: https://www.r-project.org/ (дата звернення: 14.07.2023)

24. R Markdown. URL: https://rmarkdown.rstudio.com/ (дата звернення: 14.07.2023)

25. Markdown. URL: https://uk.wikipedia.org/wiki/Markdown (дата звернення: 14.07.2023)

26. LaTeX. URL: https://uk.wikipedia.org/wiki/LaTeX (дата звернення: 14.07.2023)

27. sandino. 2013. "Cheat Sheet of Markdown." Article. https://github.com/sandino/Markdown-Cheatsheet#emphasis .

28. Pandoc a universal document converter. URL: https://pandoc.org/ (дата звернення: 14.07.2023)

29. CRAN. URL: https://cran.r-project.org/ (дата звернення: 14.07.2023)

30. Chan, Chung-hong, Geoffrey CH Chan, Thomas J. Leeper, and Jason Becker. 2018. *Rio: A Swiss-Army Knife for Data File i/o*

31. Wickham Hadley. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. http://ggplot2.org.

32. Wickham Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain Francois, Garrett Grolemund, et al. 2019. "Welcome

to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

33. Графік Q–Q. URL: https://uk.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D1%84%D1%96%D0%BA_Q-Q (дата звернення: 14.07.2023)

34. Welcome to Quarto. (n.d.). Retrieved from https://quarto.org/

35. Wickham H., Navarro D., & Pedersen T. L. (2023). *ggplot2: Elegant Graphics for Data Analysis (3e)*. Retrieved from https://ggplot2-book.org/

36. Wickham H., Çetinkaya-Rundel M., & Grolemund, G. (n.d.). *R for Data Science (2e)*. Retrieved from https://r4ds.hadley.nz/

37. Сидоренко, В. М. (2022, April 1). *Data Science на R. Лабораторний практикум (draft version)*. Retrieved from https://vgamaley.github.io/DS-book-lab/

Додаток А

Зразок оформлення титульної сторінки звіту

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КРЕМЕНЧУЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ МИХАЙЛА ОСТРОГРАДСЬКОГО

Кафедра комп'ютерної інженерії та електроніки

ПРАКТИЧНА РОБОТА
з навчальної дисципліни **«Імовірнісно-статистичні методи інформаційних технологій»**

Тема «____________________________»

Студент гр.________ПІБ_____________
Викладач _________ПІБ_____________

Кременчук 20__

Методичні вказівки щодо виконання практичних і самостійної робіт студентів з навчальної дисципліни «Імовірнісно-статистичні методи інформаційних технологій» для студентів денної форми навчання для спеціальності 123 – «Комп’ютерна інженерія» освітньо-професійної програми «Комп’ютерна інженерія» освітнього ступеня «Бакалавр» (частина 2)

Укладач доц. В. М. Сидоренко
Відповідальний за випуск д. т. н., проф. А. Л. Перекрест

Підп. до др.________________Формат 60х84 1/16. Папір тип. Друк ризографія.
Ум. друк. арк. _____Наклад ________прим. Зам. №__________ Безкоштовно.

Редакційно-видавничий відділ
Кременчуцького національного університету
імені Михайла Остроградського
вул. Університетська, 20 м. Кременчук, 39600