МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«УЖГОРОДСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ»
Факультет математики та цифрових технологій
Кафедра теорії ймовірностей і математичного аналізу

А.М. Тегза, О.О. Синявська

БІОМЕТРІЯ

МЕТОДИЧНІ ВКАЗІВКИ

до виконання типових індивідуальних завдань.

Ужгород – 2024

Біометрія: методичні вказівки до виконання типових індивідуальних завдань /Укладачі: *А.М. Тегза, О.О. Синявська.* Ужгород: ДВНЗ «УжНУ», 2024. 34 с.

У методичних вказівках наведено рекомендації для виконання індивідуальних робіт з біометрії на теми: «Ймовірність випадкової події», «Випадкові величини та їх характеристики», «Групування результатів спостережень. Обчислення статистичних показників вибірки», «Кореляційний аналіз», «Регресійний аналіз», а також завдання для самостійної роботи студентів.

Методичні вказівки розроблено для студентів спеціальності «Лісове господарство», а також студентів суміжних спеціальностей.

Рекомендовано до друку засіданням кафедри теорії ймовірностей і математичного аналізу, протокол №14 від 07 червня 2024 року.

Рекомендовано до друку науково-методичною комісією факультету математики та цифрових технологій ДВНЗ «УжНУ», протокол № 10 від 19 червня 2024 року.

Рецензенти:
докт. техн. наук, проф. Гече Ф.Є.
канд. фіз.-мат. наук, доц. Млавець Ю.Ю.

1 Теорія ймовірностей

Для написання даної лабораторної роботи потрібно повторити теоретичні питання з наступних тем:

Різні сполуки комбінаторики. Класичне, геометричне та статистичне означення ймовірності випадкової події. Основні теореми теорії ймовірностей. Послідовні незалежні випробовування, схема Бернуллі, граничні теореми. Випадкові величини. Ряд розподілу, функція розподілу. Числові характеристики випадкових величин.

Розв'язування типових задач

1. У групі з 15 бізнесменів тільки 8 мають досвід роботи у запропонованій галузі лісового господарства. Для проекту потрібно відібрати 4 особи. За припущення, що відбір претендентів проводять навмання, знайти ймовірність того, що в команду з чотирьох чоловік потраплять всі, хто має досвід роботи.

   Розв'язання. Нехай A – подія, яка полягає у відборі команди з чотирьох чоловік з досвідом роботи. Враховуючи, що розміщення всередині команди не важливе, то число способів, якими можна набрати команду 4 осіб із 15 буде рівним C_{15}^4. Аналогічно число сприятливих події A результатів експерименту буде рівним C_8^4. Тоді за класичним означенням ймовірності матимемо:

   $$P(A) = \frac{C_8^4}{C_{15}^4} = \frac{\frac{8!}{4!4!}}{\frac{15!}{11!4!}} = \frac{5 \cdot 6 \cdot 7 \cdot 8}{12 \cdot 13 \cdot 14 \cdot 15} = \frac{2}{39} \approx 0.05.$$

2. Із 20 рейсів, що виконуються з аеропорту протягом доби, 60 % рейсів виконуються на власному літаковому парку. Знайти ймовірність того, що з вибраних навмання 5 рейсів рівно 3 виконуються на власному парку.

   Розв'язання. Нехай A — подія, яка полягає в тому, що з вибраних навмання 5 рейсів рівно 3 виконуються на власному парку.

   60 % із 20 це 12, тобто 12 рейсів із 20 виконуються на власному літаковому парку, а залишкові 8 рейсів виконуються іншими літаками. Число всіх можливих навмання вибраних 5 рейсів із 20 рівне C_{20}^5. Число наборів з п'яти рейсів, які сприяють виконанню події A рівне $C_{12}^3 C_8^2$. Тоді

ймовірність настання події A рівна:

$$P(A) = \frac{C_{12}^3 C_8^2}{C_{20}^5} = \frac{\frac{12!}{3!9!}\frac{8!}{2!6!}}{\frac{20!}{5!15!}} = \frac{5 \cdot 7 \cdot 11}{3 \cdot 17 \cdot 19} \approx 0.4$$

3. Партія з 30 виробів містить 10 % браку. Знайти ймовірність того, що серед 5 виробів, узятих випадково, виявиться не більше одного бракованого.

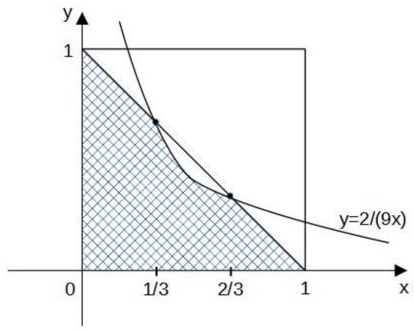
   Розв'язання. Із 30 виробів маємо 3 бракованих, 27 якісних. Нехай A – подія, яка полягає в тому, що серед 5 виробів виявиться не більше одного бракованого, тобто або 1 бракований, або всі 5 будуть якісними. Тому ймовірність цієї події (за теоремою додавання) рівна:

$$P(A) = P(0) + P(1) = \frac{C_{27}^5}{C_{30}^5} + \frac{C_3^1 C_{27}^4}{C_{30}^5} \approx 0.57 + 0.37 = 0.94$$

4. З проміжку $[0, 1]$ випадковим способом вибирають два дійсних числа. Знайти ймовірність того, що їх сума не більша одиниці, а добуток не перевищує $\frac{2}{9}$.

   Розв'язання. Нехай $x, y \in [0, 1]$. Тоді згадана умова задачі (тобто подія A) виконається, якщо x і y задовольнятимуть системі:

$$\begin{cases} x + y \leq 1, \\ xy \leq \frac{2}{9}. \end{cases}$$



Графічно розв'язок системи виглядає наступним чином:

Для обчислення ймовірності згаданої події, використаємо геометричне означення. Тобто,

$$P(A) = \frac{S_{\text{області}}}{S_{\text{квадрата}}} = \frac{\int\limits_0^{1/3} (1-x)dx + \int\limits_{1/3}^{2/3} 2/(9x)dx + \int\limits_{2/3}^{1} (1-x)dx}{1} = \frac{1}{3} + \frac{2}{9}\ln 2.$$

5. Через метеорологічні умови літак було відправлено на запасний аеродром, під час наближення до якого у баках літака залишалося палива на 3 заходи на посадку. Ймовірність посадки літака за першого заходу дорівнює 0,8, за другого – 0,95, за третього – 0,995. Знайти ймовірність вдалої посадки літака.

Розв'язання. Посадка буде вдалою, якщо виконається одна з подій A_1, A_2, A_3. Де A_1 – літак посаджено з першого разу; A_2 – літак першого разу не вдалось посадити, проте вдалось за другим заходом; A_3 – за першими двома заходами не вдалось посадити, за третім – вдалось. Тому, використовуючи теореми додавання та множення несумісних, незалежних подій, матимемо:

$$P(A) = P(A_1)+P(A_2)+P(A_3) = 0.8+0.2\cdot 0.95+0.2\cdot 0.05\cdot 0.995 = 0.99995.$$

6. Серед великої кількості дерев'яних подарункових виробів, що знаходяться в комплекті, 30 % – нестандартні. Знайти ймовірності того, що серед 5 виробів, навмання взятих із комплекту, буде: а) тільки один нестандартний; б) принаймні один нестандартний.

   Розв'язання. У даній задачі виконуються умови схеми Бернуллі, де $n = 5$, а ймовірність вибору нестандартної деталі рівна $p = 0.3$. Оскільки n невелике, а p не дуже мале, то використаємо формулу Бернуллі.

   а) $P_5(1) = C_5^1 \cdot 0.3 \cdot 0.7^4 = 0.36015$;
   б) $P_5(k \geq 1) = 1 - P_5(0) = 1 - 0.7^5 = 0.83193$.

7. Телефонна станція обслуговує 2000 абонентів. Імовірність того, що будь-який абонент зателефонує на станцію протягом години, дорівнює 0,001. Знайти ймовірності того, що протягом години на телефонну станцію зателефонують: не більше 2 абонентів.

   Розв'язання. У даній задачі виконуються умови схеми Бернуллі, але оскільки $n = 2000$ велике, а $p = 0.001$ дуже мале, то використаємо граничну теорему Пуассона:

$$P_n(k) \approx \frac{\lambda^m}{m!}e^{-\lambda}, \qquad \text{де} \quad \lambda = np = 2000 \cdot 0.001 = 2.$$

$$P_{2000}(k \leq 2) = P(0) + P(1) + P(2) = 0.1353 + 0.2707 + 0.2707 = 0.6767.$$

8. Завод виготовляє 75 % продукції вищого гатунку. Знайти ймовірність того, що із 300 виробів, випущених заводом, кількість якісних виробів буде: а) 250 виробів; б) 220 – 235.

   Розв'язання. Знову маємо виконання умов схеми Бернуллі, проте оскільки $n = 300$ досить велике, $p = 0.75$, то формулою Бернуллі не зручно користуватись, тому використаємо граничні теореми Муавра-Лапласа.

   а) За локальною теоремою Муавра-Лапласа:

$$P_n(k) \approx \frac{\varphi(x)}{\sqrt{npq}}, \qquad \text{де} \quad \varphi(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \text{ – функція Гаусса,}$$

$$x = \frac{k - np}{\sqrt{npq}} = \frac{250 - 300 \cdot 0.75}{\sqrt{300 \cdot 0.75 \cdot 0.25}} = \frac{250 - 225}{7.5} = 3.33,$$

$$P_{300}(250) = \frac{\varphi(3.33)}{7.5} = \frac{0.0016}{7.5} \approx 0.0002$$

а) За інтегральною теоремою Муавра-Лапласа:

$$P_n(k_1, k_2) \approx \Phi_0(b) - \Phi_0(a), \qquad \text{де}$$

$$a = \frac{k_1 - np}{\sqrt{npq}} = \frac{220 - 300 \cdot 0.75}{\sqrt{300 \cdot 0.75 \cdot 0.25}} = -0.67,$$

$$b = \frac{k_2 - np}{\sqrt{npq}} = \frac{235 - 300 \cdot 0.75}{\sqrt{300 \cdot 0.75 \cdot 0.25}} = 1.33.$$

Тоді $P_{300}(220, 235) \approx \Phi_0(1.33) - \Phi_0(-0.67) = 0.4082 + 0.2486 = 0.6568.$

9. Цінні папери на біржі кожного дня можуть із ймовірністю 0.5 подорожчати на 10 %. Спостереження ведеться три дні. Початкова вартість цін- них паперів 5 000 грн. Побудувати ряд розподілу випадкової величини X – вартості цінних паперів, вважаючи її розподіленою за біноміальним законом та знайти її числові характеристики.

   Розв'язання. З умови задачі вартість цінних паперів може набувати таких значень: 5000 грн. (якщо жоден з трьох днів подорожчання не відбудеться), 5500 грн. (якщо подорожчає на 10% лиш один раз), 6050 грн. (якщо два рази відбудеться подорожчання) і 6655 грн. (якщо кожного дня буде ціна збільшуватись на 10%). Ймовірності цих значень будемо обчислювати за біноміальним розподілом, тобто за формулою Бернуллі, враховуючи, що $p = 0.5$, $q = 0.5$.

   $P_3(0) = 0.5^3 = 0.125$; $\quad P_3(1) = C_3^1 \cdot 0.5 \cdot 0.5^2 = 0.375$;

   $P_3(2) = C_3^2 \cdot 0.5^2 \cdot 0.5 = 0.375$; $\quad P_3(3) = 0.5^3 = 0.125$.

   Матимемо такий ряд розподілу випадкової величини X:

X	5000	5500	6050	6655
p_i	0.125	0.375	0.375	0.125

   Тепер можемо обчислити числові характеристики випадкової величини (розрахунки проведені в електронних таблицях Excel):

$$MX = \sum_{i=1}^{4} x_i p_i = 5000 \cdot 0.125 + 5500 \cdot 0.375 + 6050 \cdot 0.375 + 6655 \cdot 0.125 =$$
$$= 5788.125.$$

$$DX = \sum_{i=1}^{4}(x_i - MX)^2 p_i = 77642.6 + 31131 + 25716,9 + 93934 = 228424.6$$

$$\sigma = \sqrt{DX} = 477.94.$$

10. У білеті є три задачі. Ймовірність правильного розв'язання студентом першої задачі дорівнює 0.9; другої – 0.8; третьої – 0.7. Побудувати ряд розподілу випадкової величини X – кількості правильно розв'язаних задач у білеті та знайти її числові характеристики.

    Розв'язання. З умови задачі випадкова величина X може прийняти значення від 0 до 3, причому ймовірності правильно виконаних номерів задач $p_1 = 0.9$, $p_2 = 0.8$, $p_3 = 0.7$. Обчислимо ймовірності значень випадкової величини:

$$\begin{aligned}
P(X=0) &= q_1 \cdot q_2 \cdot q_3 = 0.1 \cdot 0.2 \cdot 0.3 = 0.006;\\
P(X=1) &= p_1 \cdot q_2 \cdot q_3 + q_1 \cdot p_2 \cdot q_3 + q_1 \cdot q_2 \cdot p_3 =\\
&= 0.9 \cdot 0.2 \cdot 0.3 + 0.1 \cdot 0.8 \cdot 0.3 + 0.1 \cdot 0.2 \cdot 0.7 = 0.092;\\
P(X=2) &= p_1 \cdot p_2 \cdot q_3 + p_1 \cdot q_2 \cdot p_3 + q_1 \cdot p_2 \cdot p_3 =\\
&= 0.9 \cdot 0.8 \cdot 0.3 + 0.9 \cdot 0.2 \cdot 0.7 + 0.1 \cdot 0.8 \cdot 0.7 = 0.398;\\
P(X=3) &= p_1 \cdot p_2 \cdot p_3 = 0.9 \cdot 0.8 \cdot 0.7 = 0.504;
\end{aligned}$$

    Матимемо такий ряд розподілу випадкової величини X:

X	0	1	2	3
p_i	0.006	0.092	0.398	0.504

    Обчислимо числові характеристики випадкової величини (розрахунки проведені в електронних таблицях Excel):

$$MX = \sum_{i=1}^{4} x_i p_i = 0,092 + 2 \cdot 0.398 + 3 \cdot 0.504 = 2.4.$$

$$DX = \sum_{i=1}^{4}(x_i - M\xi)^2 p_i = 0.46;$$

$$\sigma = \sqrt{DX} = 0.678.$$

Завдання для самостійної роботи №1

Завдання 1

1. Деяка компанія виготовляє приймачі для геодезичних робіт. Із партії 20 приймачів, серед яких 4 несправні, для перевірки беруть 3 вироби. Яка ймовірність того, що до числа відібраних потрапить не більше одного несправного приймача?

2. Маємо два ящики. У першому 5 білих, 7 зелених кульок, у другому – 10 білих і 8 зелених. З кожного ящика навмання беруть по одній кульці. Яка ймовірність того, що обидві кульки білі?

3. Із урни, в якій 7 білих, 5 червоних і 8 зелених кульок навмання вибирають три кульки. Яка ймовірність того, що всі три кульки білі.

4. У класі 12 хлопчиків та 18 дівчаток. Треба вибрати делегацію з трьох представників цього класу. Яка ймовірність того, що у складі делегації будуть два хлопчики?

5. Із 100 лотерейних вкитікв тільки 4 виграшних. Яка ймовірність того, що два квитки виявляться виграшними?

6. Маємо два ящики. У першому 9 білих, 4 зелених кульок, у другому – 5 білих і 8 зелених. З кожного ящика навмання беруть по одній кульці. Яка ймовірність того, що кульки різних кольорів?

7. Волонтери планують посадити у міському парку 50 декоративних дерев, серед яких 10 сакур, 12 гінгобілоб, 15 каштанів Бріоті, 13 японських багряників. Кожний волонтер має посадити 5 дерев. Відомо, що Тарас перший випадковим чином вибрав 5 дерев. Яка ймовірність того, що волонтеру Тарасу попадуться 2 сакури і 3 каштани?

8. Із урни, в якій 7 білих, 5 червоних і 8 зелених кульок навмання вибирають три кульки. Яка ймовірність того, що всі три кульки червоні.

9. Деяка компанія виготовляє приймачі для геодезичних робіт. Із партії 20 приймачів, серед яких 4 несправні, для перевірки беруть 3 вироби. Яка ймовірність того, що до числа відібраних потраплять тільки справні приймачі?

10. Яка ймовірність того, що при киданні двох кубиків випаде сума цифр: 1) не менше 5; 2) більше 4?

11. У класі 10 хлопчиків та 19 дівчаток. Треба вибрати делегацію з трьох представників цього класу. Яка ймовірність того, що у складі делегації буде хоча б одна дівчинка?

12. Із урни, в якій 7 білих, 5 червоних і 8 зелених кульок навмання вибирають три кульки. Яка ймовірність того, що всі кульки різних кольорів.

13. Волонтери планують посадити у міському парку 50 декоративних дерев, серед яких 12 сакур, 11 гінгобілоб, 13 каштанів Бріоті, 14 японських багряників. Кожний волонтер має посадити 5 дерев. Відомо, що Петро перший випадковим чином вибрав 5 дерев. Яка ймовірність того, що цьому волонтеру попадуться 1 сакура, 2 японських багряника і 2 гінгобілоб?

14. У класі 11 хлопчиків та 18 дівчаток. Треба вибрати делегацію з трьох представників цього класу. Яка ймовірність того, що у складі делегації буде один хлопчик?

15. Деяка компанія виготовляє приймачі для геодезичних робіт. Із партії 20 приймачів, серед яких 4 несправні, для перевірки беруть 3 вироби. Яка ймовірність того, що до числа відібраних потраплять хоча б два справні приймачі?

16. Монету кидають три рази підряд. Яка ймовірність події B - випало рівно дві цифри;

17. У клітці сидять 6 мишей:4 чорні та 2 білі. Лаборант навмання дістає 2 мишей. Яка ймовірність того, що серед них буде одна чорна і одна біла миші?

18. На ринку продають 12 звичайних зелених смерек та 8 блакитних ялин. Для новорічних свят треба вибрати три ялини. Яка ймовірність того, що серед випадково вибраних буде хоча б одна зелена смерека?

19. Із корзини супермаркету, в якій 7 білих, 5 червоних і 3 пари зелених шкарпеток навмання вибирають три пари. Яка ймовірність того, що всі три пари будуть білими.

20. У групі із 15 студентів є 10 відмінників. За списком вибирають трьох студентів. Яка ймовірність того, що всі вибрані за списком студенти відмінники?

Завдання 2

1. У деякій родині п'ятеро дітей. Знайти ймовірність того, що серед них хоча б дві дівчинки. Ймовірність народження дівчинки рівна 0,49.

2. Ймовірність присутності студента на лекції дорівнює 0,8. Знайти ймовірність того, що із 100 студентів на лекції будуть не менше 75 та не більше 90.

3. Ймовірність виробництва бракованого виробу дорівнює 0,005. Знайти ймовірність того, що серед 1000 виробів бракованих буде не більше ніж 5.

4. Ймовірність виготовлення деякого якісного аграрного виробу дорівнює 0,9. Виготовлено 50 виробів. Чому дорівнює найімовірніше число якісних виробів та ймовірність такого числа виробів?

5. Прилад складено з 1000 незалежно працюючих елементів. Ймовірність відмови будь-якого елементу за час t дорівнює 0,002. Знайти ймовірність того, що за цей час відмовлять 3 елементи.

6. Засівний фонд має 92% насіння першого сорту. Навмання взято 150 насінин. Знайти ймовірність того, що серед них 140 будуть першого сорту.

7. Відсоток проростання пшеничного насіння становить 95%. Знайти ймовірність того, що з 2000 посіяних насінин проросте від 1880 до 1920.

8. Відділ технічного контролю перевіряє вироби на якість. Ймовірність того, що виріб бракований дорівнює 0,2. Знайти ймовірність того, що серед чотирьох перевірених виробів хоча б один виріб буде якісним.

9. У деякій родині п'ятеро дітей. Знайти ймовірність того, що серед них не більше трьох дівчаток. Ймовірність народження дівчинки рівна 0,49.

10. Ймовірність виробництва бракованого виробу дорівнює 0,008. Знайти ймовірність того, що серед 1000 виробів бракованих буде не більше ніж 4.

11. Електростанція обслуговує мережу з 10000 ламп. Ймовірність включення кожної з них 0,6. Визначити ймовірність одночасного включення від 5900 до 6100 ламп.

12. Відділ технічного контролю перевіряє вироби на якість. Ймовірність того, що виріб бракований дорівнює 0,1. Знайти ймовірність того, що серед чотирьох перевірених виробів два будуть якісними.

13. Виробництво деякої аграрної продукції дає 10% браку. Яка ймовірність того, що із 200 навмання взятих виробів 15 будуть бракованими?

14. Ймовірність присутності студента на лекції дорівнює 0,8. Знайти ймовірність того, що із 100 студентів на лекції будуть не більше 74.

15. У деякій родині п'ятеро дітей. Знайти ймовірність того, що серед них хоча б три хлопчика. Ймовірність народження хлопчика рівна 0,51.

16. Ймовірність виробництва бракованого виробу дорівнює 0,007. Знайти ймовірність того, що серед 1000 виробів бракованих буде не більше ніж 10.

17. Відділ технічного контролю перевіряє вироби на якість. Ймовірність того, що виріб бракований дорівнює 0,1. Знайти ймовірність того, що серед п'яти перевірених виробів хоча б два будуть якісними.

18. У деякій родині четверо дітей. Знайти ймовірність того, що серед них хоча б два хлопчика. Ймовірність народження дівчинки рівна 0,49.

19. Скільки разів треба кинути гральний кубик, щоб найімовірніше число появи трійки дорівнювало 35?

20. Відділ технічного контролю на заводі перевіряє авто деталі на якість. Ймовірність того, що виріб бракований дорівнює 0,05. Знайти ймовірність того, що серед чотирьох перевірених виробів два будуть якісними.

Завдання 3.

1. По мішені проведено 3 постріли. Ймовірність влучення у мішень при першому пострілі становить 0.7, при другому – 0.6, а при третьому – 0.8. Знайти ряд розподілу випадкової величини X – кількості влучень при трьох пострілах. Скласти функцію розподілу та побудувати її графік. Знайти математичне сподівання та середнє квадратичне відхилення.

2. Зенітний підрозділ відкриває вогонь по трьох бомбардувальниках. Ймовірність влучення у перший літак дорівнює 0.4, у другий – 0.5, у третій

– 0.3. Скласти ряд розподілу випадкової величини X – кількості збитих літаків, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

3. У деякому аграрному приладі працюють незалежно один від одного два блоки. Ймовірність безвідмовної роботи за деякий час першого блоку дорівнює 0.8, другого – 0.6. Нехай X – випадкова величина, яка означає кількість блоків, що працюють протягом заданого часу. Скласти ряд розподілу випадкової величини X, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

4. На складі в геодезистів є 4 діючих нівелірів та 1 непрацюючий. Навмання спеціаліст взяв 3 прилади. Скласти ряд розподілу випадкової величини X, яка означає кількість діючих нівелірів серед трьох взятих. Знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

5. У ящику 5 білих та 3 чорних кульки. Навмання виймають 2 кульки. Нехай X – випадкова величина, яка означає кількість чорних кульок серед вилучених. Скласти ряд розподілу випадкової величини X, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

6. Є 7 світлодіодних ламп, з яких три не працюють. Навмання беруть 4 лампи. Випадкова величина X означає кількість працюючих ламп серед вибраних. Записати ряд та функцію розподілу величини X. Знайти математичне сподівання та середнє квадратичне відхилення.

7. По мішені проведено 3 постріли. Ймовірність влучення у мішень при першому пострілі становить 0.3, при другому – 0.2, а при третьому – 0.7. Знайти ряд розподілу кількості влучень при трьох пострілах. Скласти функцію розподілу та побудувати її графік. Знайти математичне сподівання та середнє квадратичне відхилення кількості влучень.

8. ППО відкриває вогонь по трьох ракетах ворога. Ймовірність влучення у першу ракету дорівнює 0.7, у другу – 0.8, у третю – 0.5. Скласти ряд розподілу випадкової величини X – кількості збитих ракет, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію величини.

9. Баскетболіст кидає м'яч в кільце тричі. Ймовірність влучення в кільце при одному кидку – 0.8. Випадкова величина X – число промахів баскетболіста в трьох кидках. Записати ряд та функцію розподілу величини X. Знайти математичне сподівання та середнє квадратичне відхилення.

10. У коробці 5 білих, 2 червоних та 1 синя крейд. Лектор навмання виймає 2 крейди. Нехай X – випадкова величина, яка означає кількість червоних крейд серед вибраних. Скласти ряд розподілу випадкової величини X, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

11. У партії з 30 деталей 5 бракованих. Перевіряють навмання вибрані 3 деталі. Скласти ряд розподілу випадкової величини X – число бракованих деталей серед 3 вибраних. Знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

12. Футболіст намагається забити м'яч у ворота тричі. Ймовірність влучення у ворота при одному ударі – 0.7. Випадкова величина X – число промахів футболіста в трьох ударах. Записати ряд та функцію розподілу величини X. Знайти математичне сподівання та середнє квадратичне відхилення.

13. ППО відкриває вогонь по трьох ракетах ворога. Ймовірність влучення у першу ракету дорівнює 0.6, у другу – 0.7, у третю – 0.5. Скласти ряд розподілу випадкової величини X – кількості збитих ракет, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію величини.

14. Нехай випадкова величина X – число випадань герба у 5 незалежних підкиданнях монети. Записати ряд та функцію розподілу величини ξ. Знайти математичне сподівання та середнє квадратичне відхилення.

15. На полиці 6 червоних та 3 зелених зошити. Навмання взяли 2 зошити. Нехай X – випадкова величина, яка означає кількість червоних зошитів серед вилучених. Скласти ряд розподілу випадкової величини X, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

16. Зенітний підрозділ відкриває вогонь по трьох бомбардувальниках. Ймовірність влучення у перший літак дорівнює 0.5, у другий – 0.6, у третій – 0.7. Скласти ряд розподілу випадкової величини X – кількості збитих

літаків, знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

17. На складі в геодезистів є 5 діючих тахеометрів та 2 непрацюючих. Навмання спеціаліст взяв 3 прилади. Скласти ряд розподілу випадкової величини X, яка означає кількість діючих тахеометрів серед трьох взятих. Знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

18. Нехай випадкова величина X – число випадань решки при 4 незалежних підкиданнях монети. Записати ряд та функцію розподілу величини X. Знайти математичне сподівання та середнє квадратичне відхилення.

19. Снайпер-початківець проводить 3 постріли. Ймовірність влучення у мішень при першому пострілі становить 0.5, при другому – 0.7, а при третьому – 0.8. Знайти ряд розподілу кількості влучень при трьох пострілах. Скласти функцію розподілу та побудувати її графік. Знайти математичне сподівання та середнє квадратичне відхилення кількості влучень.

20. У партії з 35 китайських мікросхем 7 бракованих. Перевіряють навмання вибрані 3 схеми. Скласти ряд розподілу випадкової величини X – число бракованих мікросхем серед 3 вибраних. Знайти функцію розподілу та побудувати її графік. Обчислити математичне сподівання та дисперсію.

2 Математична статистика (частина 1).

Для написання даної лабораторної роботи потрібно повторити теоретичні питання з наступних тем:

Статистичний розподіл вибірки. Графічне її представлення. Емпірична функція розподілу. Вибіркові характеристики.

Розв'язування типових задач

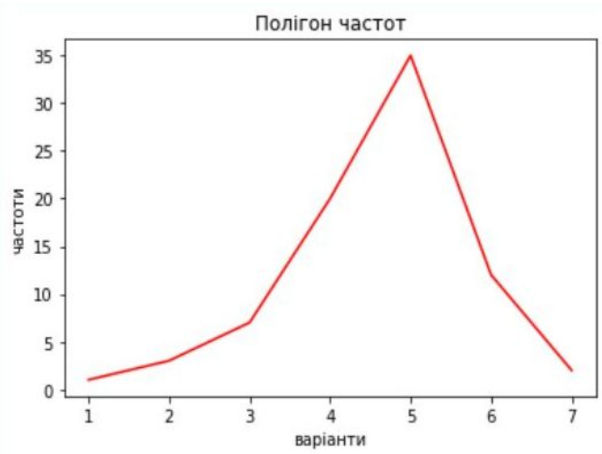
1. Вчені дослідили, що чистопородисті кішки рідко народжують більше 5 кошенят, тоді як непородисті – 5-10 кошенят. Дослідили 80 породистих кішок та їх потомство. Одержали такі результати кількості народжених кошенят у 80 кішок: 6, 2, 5, 5, 3, 5, 5, 4, 6, 5, 5, 4, 2, 6, 5, 5, 3, 5, 5, 5, 7, 5, 5, 5, 5, 3, 7, 4, 6, 4, 4, 5, 5, 6, 6, 4, 5, 4, 6, 4, 5, 4, 5, 5, 4, 5, 4, 3, 5, 4, 5, 6, 1, 6, 4, 4, 4, 5, 5, 3, 5, 5, 4, 6, 4, 6, 2, 3, 4, 5, 5, 5, 5, 5, 4, 5, 5, 6, 4, 3. Побудувати статистичний розподіл частот вибірки, зобразити полігон частот, побудувати емпіричну функцію розподілу. Обчислити основні вибіркові характеристики.

   Розв'язання. Підрахувавши скільки разів кожна варіанта зустрілась у вибірці одержимо її статистичний розподіл:

x_i	1	2	3	4	5	6	7
n_i	1	3	7	20	35	12	2

   Побудуємо полігон частот та емпіричну функцію розподілу за формулою: $F_n^*(x) = \frac{1}{n} \sum\limits_{n_i:\, x_i < x} n_i$.

$$F_n^*(x) = \begin{cases} 0, & x \le 1, \\ \frac{1}{80}, & 1 < x \le 2, \\ \frac{4}{80}, & 2 < x \le 3, \\ \frac{11}{80}, & 3 < x \le 4, \\ \frac{31}{80}, & 4 < x \le 5, \\ \frac{66}{80}, & 5 < x \le 6, \\ \frac{78}{80}, & 6 < x \le 7, \\ 1, & x > 7. \end{cases}$$



   Обчислимо основні вибіркові характеристики:

Мода: $Mo = 5$, медіана: $Me = 5$,
вибіркове середнє: $x_{\text{в}} = \frac{1}{n}\sum_{i=1}^{n} x_i n_i = 4.61$,
вибіркова дисперсія: $D_{\text{в}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - x_{\text{в}})^2 n_i = 1.24$,
вибіркове середньо-квадратичне відхилення $\sigma_{\text{в}} = 1.11$.

2. Для заданої вибіркової сукупності

$$X = [3., 2.9, 3.3, 2.7, 2.9, 2.8, 3.1, 3.1, 3.3, 3., 3., 3.1, 3., 2.8, 3.2, 2.8, 3.1, 3.5, 3.3, 2.9, 3.1, 3.1, 2.7, 3.3, 2.9, 2.8, 2.8, 3.1, 3.3, 3., 2.8, 3.2, 3.1, 2.9, 3., 3., 3.2, 2.9, 2.9, 3.2, 2.9, 3., 2.9, 3.2, 3.1, 3.4, 2.9, 3.2, 3.2, 3., 3.1, 2.9, 3.3, 3.2, 3.2, 3.3, 2.7, 3.1, 2.5, 2.9, 3., 3.2, 2.9, 2.7, 2.8, 2.9, 2.8, 3., 2.8, 2.8]$$

скласти інтервальний розподіл частот та відносних частот. Побудувати гістограми частот та відносних частот.

Розв'язання. Складемо інтервальний розподіл частот та відносних частот. Згідно теоретичного курсу кількість інтервалів рівне: $k = \sqrt{n} = \sqrt{70} \approx 8$, а довжина інтервалу: $h = \frac{x_{max} - x_{min}}{k} = \frac{3.5 - 2.5}{8} \approx 0.13$. Тоді весь варіаційний проміжок $[2.5, 3.5]$ ділимо на такі інтервали: $[2.5, 2.63)$, $[2.63, 2.76)$, $[2.76, 2.89)$, $[2.89, 3.02)$, $[3.02, 3.15)$, $[3.15, 3.28)$, $[3.28, 3.41)$, $[3.41, 3.5]$. Підрахувавши скільки варіант попадають у кожний інтервал, складемо інтервальний статистичний розподіл частот і відносних частот:

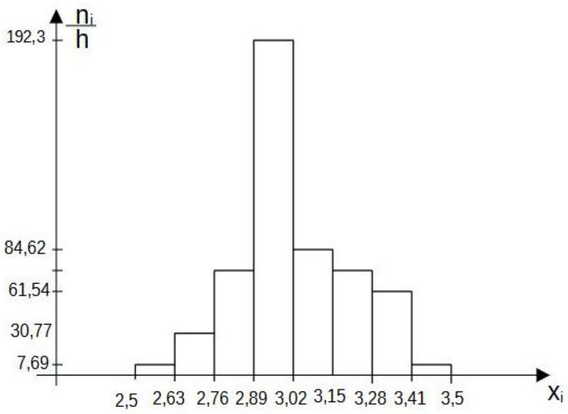
(x_i, x_{i+1})	$[2.5,\ 2.63)$	$[2.63,\ 2.76)$	$[2.76,\ 2.89)$	$[2.89,\ 3.02)$
n_i	1	4	10	25
$w_i = \frac{n_i}{n}$	$\frac{1}{70} \approx 0.01$	$\frac{4}{70} \approx 0.06$	$\frac{10}{70} \approx 0.01$	$\frac{25}{70} \approx 0.36$
	$[3.02,\ 3.15)$	$[3.15,\ 3.28)$	$[3.28,\ 3.41)$	$[3.41,\ 3.5]$
n_i	11	10	8	1
$w_i = \frac{n_i}{n}$	$\frac{11}{70} \approx 0.16$	$\frac{10}{70} \approx 0.01$	$\frac{8}{70} \approx 0.01$	$\frac{1}{70} \approx 0.01$

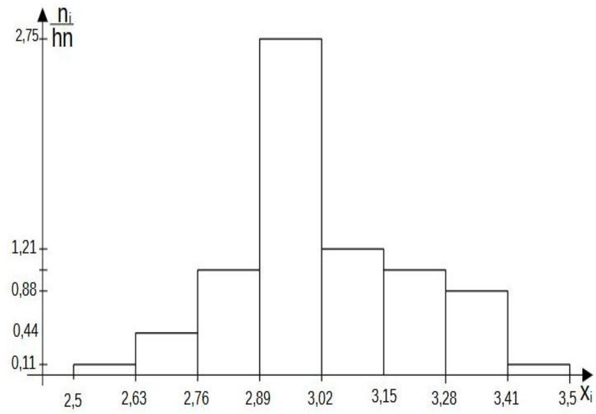
Гістограма частот являє собою фігуру, яка складається з прямокутників, кожний з яких має основу h і висоту $\frac{n_i}{h}$.

Гістограма відносних частот є фігурою, що складається з прямокутників, кожний з яких має основу завдовжки h і висоту $\frac{w_i}{h}$. Тому для побудови гістограм випишемо таблицю інтервалів із їхніми значеннями $\frac{n_i}{h}$ і $\frac{w_i}{h}$.

(x_i, x_{i+1})	$[2.5,\ 2.63)$	$[2.63,\ 2.76)$	$[2.76,\ 2.89)$	$[2.89,\ 3.02)$
$\frac{n_i}{h}$	7.69	30.77	76.92	192.3
$\frac{w_i}{h}$	0.11	0.44	1.1	2.75
	$[3.02,\ 3.15)$	$[3.15,\ 3.28)$	$[3.28,\ 3.41)$	$[3.41,\ 3.5]$
$\frac{n_i}{h}$	84.62	76.92	61.54	7.69
$\frac{w_i}{h}$	1.21	1.1	0.88	0.11

Тепер будуємо гістограми частот і відносних частот.





Завдання для самостійної роботи №2

Завдання 1. Задано деяку вибіркову сукупність X. Скласти дискретний статистичний розподіл частот вибірки та інтервальний розподіл частот, зобразити полігон та гістограму частот, побудувати емпіричну функцію розподілу. Обчислити основні вибіркові характеристики.

1) X=[1, 0, 3, 0, 1, 2, 0, 0, 1, 3, 1, 3, 0, 3, 3, 3, 3, 0, 2, 2, 2, 1, 3, 1, 5, 2, 0, 3, 3, 1, 0, 2, 1, 3, 3, 0, 1, 1, 3, 4, 5, 1, 3, 0, 3, 0, 5, 2, 2, 3]

2) X=[0.3, -1.5, -1.9, 1.5, -1.5, 0.3, -1.5, -1.9, 1.5, -1.9, 1.5, 0.3, 1.5, -1.5, -1.5, -1.9, 0.3, 0.8, 0.3, -1.5, 0.8, -1.5, 0.8, -1.5, 0.3, 0.3, 0.8, -1.5, 1.5, 0.8, 0.8, 0.8, 1.5, -1.9, 0.3, 1.5, -1.5, 0.3, 0.8, -1.5, -1.5, 0.3, 0.8, 1.5, 0.3, 0.3, 1.5, 0.3, -1.9, 0.3]

3) X=[6, 6, 6, 6, 5, 6, 4, 5, 4, 4, 6, 6, 4, 4, 6, 5, 4, 6, 6, 5, 5, 6, 6, 4, 5, 6, 6, 4, 5, 6, 6, 4, 5, 5, 5, 5, 4, 5, 6, 4, 5, 4, 4, 6, 4, 6, 4, 6, 5, 4]

4) X=[0.4, 0.1, 0.8, 0.9, 0.6, 0.3, 0.6, 0.8 , 0.1, 0.3, 0.3, 0.2, 0.9, 0.8, 0.1 , 0.1, 0.4, 0.3, 0.2, 0.8, 0.8, 0.9, 0.6, 0.6, 0.2, 0.9, 0.8, 0.1 , 0.3, 0.9, 0.6, 0.4, 0.4, 0.8, 0.8, 0.3, 0.9, 0.6, 0.6, 0.8, 0.2, 0.1, 0.3, 0.4, 0.4, 0.4, 0.9, 0.3, 0.3, 0.2]

5) X=[0.2, 0.1, 0. , 0.2, -0.1, 0.1, 0. , 0.1, -0.1, 0.2, -0.2, 0.2, 0.1, -0.1, -0.1, -0.2, -0.1, -0.2, 0.2, 0.1, 0.1, 0. , 0.1, -0. , 0.1, 0.2, 0.2, -0.1, 0. , -0. , 0.1, -0.1, -0.2, -0.1, -0. , -0.1, -0.2, -0.1, -0.1, -0.1, -0. , 0.2, -0.2, -0.2, -0.2, 0.1, -0.1, 0.1, 0.1, 0.1, -0.1, -0.1, 0.2, -0.1, -0.2, 0.2, 0.1, -0.1, -0. , 0.1]

6) X=[5, 5, 7, 5, 5, 6, 5, 7, 7, 6, 7, 6, 5, 6, 5, 5, 5, 5, 6, 7, 6, 6, 7, 6, 5, 7, 6, 6, 7, 7, 7, 6, 6, 5, 7, 7, 6, 7, 7, 5, 5, 7, 6, 5, 7, 5, 6, 7, 5, 6, 6, 6, 6, 7, 6, 5, 6, 5, 6, 5]

7) X=[0.1, 0.1, 0.5, 0.1, 0.1, 0.2, 0.1, 0.2, 0.4, 0.1, 0.1, 0.3, 0.4, 0.5, 0.2, 0.4, 0.4, 0.1, 0.5, 0.4, 0.2, 0.4, 0.4, 0.4, 0.1, 0.3, 0.2, 0.3, 0.1, 0.2, 0.3, 0.4, 0.2, 0.3, 0.3, 0.3, 0.2, 0.3, 0.1, 0.4, 0.3, 0.4, 0.4, 0.1, 0.3, 0.4, 0.2, 0.2, 0.2, 0.4, 0.2, 0.3, 0.1, 0.4, 0.4]

8) X=[1.2, 1.4, 1.4, 1.2, 1.2, 1.5, 1.3, 1.4, 1.2, 1.2, 1.4, 1.3, 1.2, 1.5, 1.1, 1.4, 1.2, 1.3, 1.2, 1.4, 1.3, 1.4, 1.2, 1.5, 1.4, 1.2, 1.1, 1.1, 1.3, 1.4, 1.3, 1.2, 1.1, 1.2, 1.2, 1.2, 1.2, 1.2, 1.3, 1.5, 1.2, 1.3, 1.2, 1.5, 1.3, 1.3, 1.2, 1.2, 1.4, 1.3, 1.3, 1.1, 1.4, 1.2, 1.2, 1.5, 1.3, 1.3, 1.3, 1.3, 1.1, 1.2, 1.5, 1.4, 1.2]

9) [3, 2, 2, 2, 2, 5, 5, 5, 5, 2, 5, 5, 4, 3, 5, 2, 3, 3, 5, 5, 4, 4, 3, 3, 5, 5, 2, 5, 5, 5, 2, 4, 4, 3, 3, 5, 5, 5, 4, 3, 4, 5, 5, 4, 2, 3, 2, 3, 4, 3, 2, 3, 4, 3, 3, 4, 3, 5, 4, 3, 2, 4, 3, 2, 3]

10) X=[3.3, 3.3, 3.4, 3.3, 3.4, 3.5, 3.3, 3.4, 3.4, 3.3, 3.5, 3.2, 3.5, 3.6, 3.2, 3.4, 3.5, 3.5, 3.4, 3.5, 3.3, 3.2, 3.5, 3.4, 3.5, 3.2, 3.4, 3.5, 3.3, 3.4, 3.5, 3.2, 3.5, 3.6, 3.5, 3.2, 3.6, 3.4, 3.5, 3.3, 3.6, 3.3, 3.3, 3.4, 3.2, 3.4, 3.3, 3.4, 3.3, 3.3, 3.6, 3.6, 3.3, 3.2, 3.5, 3.3, 3.2, 3.4, 3.6, 3.4, 3.4, 3.3, 3.4, 3.4, 3.3]

11) X=[6, 5, 8, 5, 5, 5, 5, 7, 7, 7, 5, 6, 7, 6, 7, 6, 7, 7, 7, 5, 7, 5, 6, 7, 6, 5, 6, 6, 6, 7, 5, 5, 7, 6, 7, 8, 6, 6, 7, 7, 6, 8, 7, 5, 6, 6, 5, 6, 8, 5]

12) X=[-0.3, 0. , 0.2, 0.1, 0.2, 0.2, 0.4, 0.2, 0.3, 0.4, 0.3, 0. , -0.2, 0.1, 0.3, 0. , -0.1, 0. , 0. , 0.2, 0.1, -0.3, -0.2, 0. , -0.2, -0.5, 0.1, -0.1, 0.1, 0.5, 0.1, -0.4, 0.1, -0.2, -0.1, -0.2, 0.1, -0.3, -0.3, 0. , 0. , 0.1, -0.1, 0.5, 0. , -0.1, 0.1, 0. , -0.2, -0.1]

13) X=[5.5, 5.2, 5.2, 5.3, 5. , 5.5, 5.3, 5.5, 5.2, 5. , 5.2, 5.3, 5.2, 5. , 5.2, 5.2, 5.1, 5.1, 5.3, 5.3, 5.3, 5.4, 5.4, 5.2, 5.4, 5.3, 5.4, 5.2, 5.4, 5.1, 5.4, 5.2, 5.1, 5.4, 5.5, 5.3, 5.3, 5.3, 5.1, 5.2, 5. , 5.4, 5.3, 5.1, 5.5, 5.4, 5.5, 5.2, 5.3, 5.]

14) X=[5, 7, 7, 6, 7, 5, 6, 6, 5, 6, 6, 5, 7, 5, 6, 7, 5, 5, 7, 6, 7, 5, 7, 6, 7, 5, 5, 7, 7, 7, 5, 7, 5, 6, 6, 5, 5, 5, 5, 7, 5, 5, 7, 7, 6, 7, 7, 6, 6, 6]

15) X=[1. , 1.3, 1. , 1.1, 1. , 1. , 0.8, 0.9, 1.3, 0.8, 0.9, 1.2, 1.2, 1. , 0.7, 1.2, 0.9, 1.2, 1. , 1.1, 0.9, 1.1, 1.1, 0.8, 0.9, 1.3, 0.7, 0.7, 1.1, 0.8, 1. , 0.6, 0.9, 1. , 1. , 1.3, 1. , 1.2, 1.3, 1.1, 1. , 0.8, 1. , 0.9, 0.9, 1.2, 1.5, 1.4, 1.1, 0.9, 1.1, 0.7, 1. , 0.9, 1.1]

16) X=[1, 0, 0, 1, 0, 1, 2, -1, 0, 1, -1, 1, 0, 1, 2, 2, -1, 1, 0, 1, 1, 2, 2, 1, 2, 1, 0, -1, 1, 0, 1, -1, 1, -1, 1, 0, 2, 2, -2, 1, 1, 0, 2, 0, 0, 0, 1, 1, 1, 3, 2, -1, 1, 0, 0]

17) X= [5. , 5. , 5. , 5. , 5. , 4.9, 4.9, 5.1, 4.9, 5.1, 5. , 5.2, 4.9, 4.9, 5.1, 5.1, 5.1, 5.2, 5.1, 5. , 5.1, 5.1, 5. , 5. , 5. , 5. , 5.1, 5. , 5. , 5.1, 4.9, 5. , 5.1, 5. , 5.1, 5.1, 5.1, 4.9, 4.9, 5.1, 4.9, 5. , 4.9, 4.9, 4.7, 5. , 5. , 5. , 5.1, 5. , 5. , 5. , 5.1, 5.1, 4.9]

18) X=[9, 9, 10, 8, 11, 11, 10, 7, 10, 8, 9, 8, 7, 11, 10, 7, 9, 11, 8, 8, 10, 7, 11, 11, 11, 10, 11, 8, 8, 7, 9, 8, 8, 9, 10, 10, 9, 11, 8, 9, 8, 7, 9, 11, 8, 8, 9, 10, 11, 9, 11, 9, 9, 10, 8]

19) X=[-2. , -2. , -1.8, -1.9, -2.1, -1.9, -2.1, -1.8, -2.1, -2.1, -1.9, -2. , -2. , -2. , -2.1, -2. , -1.9, -2. , -2.1, -1.8, -1.8, -2.2, -2.1, -2. , -1.9, -1.7, -2.1, -1.9, -1.9, -2. , -1.8, -1.9, -1.9, -2.1, -2.1, -2. , -2.1, -2.1, -2. , -2. , -1.8, -2. , -1.9, -2. , -2.2, -2.1, -2.2, -1.9, -1.9, -2. , -2.2, -2.1, -1.9, -2. , -2.1]

20) X=[4, 4, 6, 5, 6, 5, 5, 4, 4, 5, 5, 5, 6, 4, 4, 6, 4, 4, 6, 5, 5, 6, 4, 4, 4, 5, 4, 4, 4, 4, 5, 6, 6, 5, 6, 5, 5, 6, 6, 5, 4, 5, 6, 4, 5, 6, 5, 4, 5, 4, 5, 6, 6, 5, 5]

Завдання 2. Заміряйте довжину листків дубу звичайного (або іншого дерева, 50-60 од.), запишіть отримані дані. З отриманих замірів утворіть ранжирований варіаційний ряд. Охарактеризуйте як варіює досліджувана ознака. Знайдіть вибіркове середнє, дисперсію, стандартне відхилення, коефіцієнт варіації.

Завдання 3. Виконати засобами електронних таблиць Libre Office Calc або MS Excel.

1) Для кожного району Київської області обчислити сумарну чисельність населення відповідного району, середнє значення населення району і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

2) Для кожного району Тернопільської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

3) Для кожного району Чернігівської області обчислити сумарну чисельність населення відповідного району, середнє значення населення району і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

4) Для кожного району Київської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

5) Для кожного району Черкаської області обчислити сумарну чисельність населення відповідного району, середнє значення населення району і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

6) Користуючись даними, що містяться у файлі `екологічна_радіаційна обстановка.xlsx` обчислити середнє середньодобове значення радіоактивності газо-аерозольних викидів інертних радіоактивних газів у

навколишнє середовище (стовпець irg) за кожний рік окремо і середнє значення і виправлене стандартне відхилення викидів за 5 років (2018-2022рр) для всіх АЕС. Результати подати на діаграмах, проаналізувати і зробити висновки.

7) Для кожного району Закарпатської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

8) Користуючись даними, що містяться у файлі `екологічна_радіаційна обстановка.xlsx` обчислити середнє середньодобове значення радіоактивності газо-аерозольних викидів радіонуклідів йоду у навколишнє середовище (стовпець `iodine_radionuclides`) за кожний рік окремо і середнє значення і виправлене стандартне відхилення викидів за 5 років (2018-2022рр) для всіх АЕС. Результати подати на діаграмах, проаналізувати і зробити висновки.

9) Для кожного району Вінницької області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

10) Для кожного району Рівненської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

11) Для кожного району Волинської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

12) Для кожного району Полтавської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

13) Для кожного району Житомирської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

14) Для кожного району Сумської області обчислити сумарну чисельність населення відповідного району, середнє значення населення району і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

15) Для кожного району Івано-Франківської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

16) Користуючись даними, що містяться у файлі `екологічна_радіаційна обстановка.xlsx` обчислити середній об'єм водних скидів за звітний період (стовпець `volume`) за кожний рік окремо і середнє значення і виправлене стандартне відхилення викидів за 5 років (2018-2022рр) для всіх АЕС. Результати подати на діаграмах, проаналізувати і зробити висновки.

17) Для кожного району Львівської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

18) Для кожного району Закарпатської області обчислити сумарну чисельність населення відповідного району, середнє значення населення району і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

19) Для кожного району Миколаївської області обчислити сумарну земельну площу населених пунктів відповідного району, середнє значення площі і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

20) Для кожного району Тернопільської області обчислити сумарну чисельність населення відповідного району, середнє значення населення району і його середнє квадратичне відхилення. Всі результати показати графічно, проаналізувати і зробити висновки. Дані містяться у файлі `dovidnik-pokaznikiv-ngo-np-01-04-2023.xlsx`.

3 Математична статистика (частина 2).

Для написання даної лабораторної роботи потрібно повторити теоретичні питання з наступних тем:

Кореляційний аналіз. Побудова парної лінійної регресійної моделі. Критерії для оцінки моделі на точність. Побудова оптимальної нелінійної регресійної моделі, оцінка її на точність.

Розв'язування типових задач

1. За даними про результати замірів діаметрів X і висот Y дерев у ялицевому насадженні потрібно обчислити коефіцієнт кореляції між діаметрами ($\mathrm{x_i}$) і висотами (y_i) 20 дерев, дати його оцінку.

$\mathrm{X} = x_i$	12,9	16,0	17,8	24,2	19,6	17,5	16,8	16,1	17,4	16,6
$Y = y_i$	16,3	14,6	20,1	22,7	18,7	19,0	18,8	19,9	20,3	19,6

$\mathrm{X} = x_i$	19,1	13,5	11,3	25,1	13,1	26,4	16,7	13,9	30,7	23,7
$Y = y_i$	20,4	17,6	12,9	21,4	15,9	22,2	19,6	14,4	22,4	22,5

Знайти рівняння прямої регресії, оцінити значущість коефіцієнтів регресії прямої.

Розв'язання. **1.** Для оцінки тісноти зв'язку між X і Y розрахуємо коефіцієнт кореляції (Пірсона) за формулою:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

Розрахунки для зручності проведемо у вигляді таблиці.

i	x_i	y_i	$x_i-\overline{x}$	$y_i-\overline{y}$	$(x_i-\overline{x})^2$	$(y_i-\overline{y})^2$	$(x_i-\overline{x})(y_i-\overline{y})$
1	12,9	16,3	-5,52	-2,66	30,470	7,076	14,683
2	16,0	14,6	-2,42	-4,36	5,876	19,010	10,551
3	17,8	20,1	-0,62	1,14	0,384	1,30	-0,707
4	24,2	22,7	5,78	3,74	33,408	13,988	21,617
5	19,6	18,7	1,18	-0,26	1,392	0,068	-0,307
6	17,5	19,0	-0,92	0,04	0,846	0,002	-0,037
7	16,8	18,8	-1,62	-0,16	2,624	0,026	0,259
8	16,1	19,9	-2,32	0,94	5,382	0,884	-2,181
9	17,4	20,3	-1,02	1,34	1,040	1,796	-1,367
10	16,6	19,6	-1,82	0,64	3,312	0,410	-1,165
11	19,1	20,4	0,68	1,44	0,462	2,074	0,979
12	13,5	17,6	-4,92	-1,36	24,206	1,850	6,691
13	11,3	12,9	-7,12	-6,06	50,694	36,724	43,147
14	25,1	21,4	6,68	2,44	44,622	5,954	16,299
15	13,1	15,9	-5,32	-3,06	28,302	9,364	16,279
16	26,4	22,2	7,98	3,24	63,680	10,498	25,855
17	16,7	19,6	-1,72	0,64	2,958	0,410	-1,101
18	13,9	14,4	-4,52	-4,56	20,430	20,794	20,611
19	30,7	22,4	12,28	3,44	150,798	11,834	42,243
20	23,7	22,5	5,28	3,54	27,878	12,532	18,691
Σ	368,4	379,3	0	0	498,744	156,594	231,040

Отже,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{368,4}{20} = 18,42; \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{379,3}{20} = 18,96,$$

і

$$r = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}} = \frac{231,040}{\sqrt{498,744\cdot 156,594}} = 0,827;$$

$$r^2 = (0,827)^2\cdot 100\% = 68\%.$$

За значенням коефіцієнта кореляції можна зробити висновок, що між X і Y існує сильний додатній зв'язок.

Перевіримо статистичну значущість знайденого коефіцієнта кореляції. Знайдемо оцінку значущості коефіцієнта кореляції за Z–перетворенням Фішера:

$$z = 0,5\ln\frac{1+r}{1-r} = 0,5\ln\frac{1+0,827}{1-0,827} = 1,178.$$

Основна помилка Z–перетворення Фішера становить:

$$m_z = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{20-3}} = 0,242.$$

Фактичне значення t–критерію Ст'юдента дорівнює:

$$t = \frac{z}{m_z} = \frac{1,178}{0,242} = 4,87.$$

Із таблиці значень t–критерію Ст'юдента за числом ступенів свободи $l = n - 2 = 20–2 = 18$ знайдемо критичне значення: $t_{5\%} = 2,10$.

Виходячи з того, що фактичне значення t–критерію Ст'юдента перевищує критичне $(4,87 > 2,10)$, можна стверджувати, що обчислений коефіцієнт кореляції є значущим показником зв'язку між досліджуваними ознаками на рівні ймовірності 95%.

2. Коефіцієнт рангової кореляції Спірмена є непараметричним показником кореляційного зв'язку і застосовується при вивченні кількісних та якісних ознак, якщо наперед невідомі закон розподілу та форма зв'язку.

Варіанти за діаметром розмістимо у порядку зростання і запишемо відповідні їм висоти. Значенням діаметрів і висот присвоюють відповідні порядкові номери, або ранги (X, Y). У випадку однакових значень знайдемо середні арифметичні значення з їх порядкових номерів. Для обчислення коефіцієнта кореляції обчислюємо квадрати різниць рангів d_i^2, де d_i — різниці між рангами i–го значення X та відповідного значення Y. Порядок обчислень наведено в наступній таблиці.

x_i	y_i	Ранг X	Ранг Y	d_i	d_i^2
11,3	12,9	1	1	0	0
12,9	16,3	2	5	-3	9
13,1	15,9	3	4	-1	1
13,5	17,6	4	6	-2	4
13,9	14,4	5	2	3	9
16,0	14,6	6	3	3	9
16,1	19,9	7	12	-5	25
16,6	19,6	8	10,5	-2,5	6,25
16,7	19,6	9	10,5	-1,5	2,25
16,8	18,8	10	8	2	4
17,4	20,3	11	14	-3	9
17,5	19,0	12	9	3	9
17,8	20,1	13	13	0	0
19,1	20,4	14	15	-1	1
19,6	18,7	15	7	8	64
23,7	22,5	16	19	-3	9
24,2	22,7	17	20	-3	9
25,1	21,4	18	16	2	4
26,4	22,2	19	17	2	4
30,7	22,4	20	18	2	4
368,4	379,3	210	210	0	182,5

Здійснюємо контроль обчислень:

$$\frac{n(n+1)}{2} = \sum X = \sum Y = \frac{20(20+1)}{2} = 210.$$

Коефіцієнт кореляції дорівнює:

$$r_s = 1 - \frac{6\sum d_i^2}{n^3 - n} = 1 - \frac{6 \cdot 182,5}{20^3 - 20} = 0,863.$$

Оцінку значущості коефіцієнта рангової кореляції проведемо шляхом порівняння його фактичного і критичного значення. Із таблиці критичних значень коефіцієнта кореляції (r_{st}) залежно від обсягу вибірки на рівнях значущості 5 % і 1 %. Якщо $r_s \geq r_{st}$, коефіцієнт кореляції вважається значущим показником зв'язку, а якщо $r_s < r_{st}$ — незначущим.

За таблицею критичних значень при $n = 20$, $r_{5\%} = 0,534$.

Оскільки фактичне значення коефіцієнта рангової кореляції Спірмена перевищує критичне значення ($0,863 > 0,534$), кореляційний зв'язок між діаметрами і висотами дерев ялиці є значущим на рівні ймовірності 95 %.

3. При проведенні кореляційного аналізу виявлено високий і прямолінійний зв'язок між діаметрами і висотами дерев у ялицевому насадженні, що дозволяє побудувати рівняння регресії. Формуємо таблицю допоміжних значень для підрахунку сум, необхідних для визначення коефіцієнтів регресії.

i	x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
1	12,9	16,3	166,41	265,69	210,27
2	16,0	14,6	256	213,16	233,6
3	17,8	20,1	316,84	404,01	357,78
4	24,2	22,7	585,64	515,29	549,34
5	19,6	18,7	384,16	349,69	366,52
6	17,5	19,0	306,25	361	332,5
7	16,8	18,8	282,24	353,44	315,84
8	16,1	19,9	259,21	396,01	320,39
9	17,4	20,3	302,76	412,09	353,22
10	16,6	19,6	275,56	384,16	325,36
11	19,1	20,4	364,81	416,16	389,64
12	13,5	17,6	182,25	309,76	237,6
13	11,3	12,9	127,69	166,41	145,77
14	25,1	21,4	630,01	457,96	537,14
15	13,1	15,9	171,61	252,81	208,29
16	26,4	22,2	696,96	492,84	586,08
17	16,7	19,6	278,89	384,16	327,32
18	13,9	14,4	193,21	207,36	200,16
19	30,7	22,4	942,49	501,76	687,68
20	23,7	22,5	561,69	506,25	533,25
Σ	368,4	379,3	7284,68	7350,01	7217,75

Визначаємо параметри a та b із рівняння лінійної регресії $\overline{y}_x = a + b \cdot x$ за формулами:

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} =$$

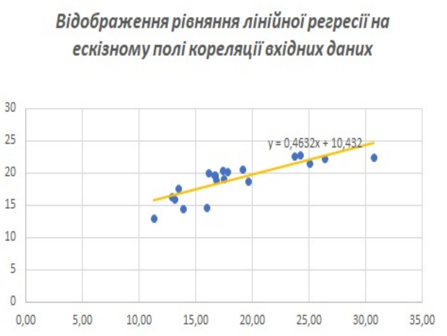
$$= \frac{379,30 \cdot 7284,68 - 368,40 \cdot 7217,75^2}{20 \cdot 7284,62 - (368,40)} = 10,432;$$

$$b = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} =$$

$$= \frac{20 \cdot 7217,75 - 368,40 \cdot 379,30^2}{20 \cdot 7284,62 - (368,40)} = 0,463.$$

Підставляємо отримані параметри до загального рівняння лінійної регресії та отримаємо результат:

$$\overline{y}_x = a + b \cdot x = 10,432 + 0,463x.$$



Завдання для самостійної роботи №3

Завдання 1. Дослідити кореляційну залежність між двома ознаками X і Y, обчисливши коефіцієнт кореляції r. Перевірити гіпотезу про значущість r і знайти похибку обчислень.

1) Набір даних `used_cars_UA.csv` (для автомобілей на бензині); X= пробіг (у милях), Y=ціна (у фунтах)

2) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); X_1 = women 16-24 , X_2 = men 16-24.

3) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); X_1 = women 25-54 , X_2 = men 25-54.

4) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); X_1 = women 55-64, X_2 = men 55-64.

5) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); X_1 = `women_high_school`, X_2 = `men_high_school`.

6) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); X_1 = `women_some_college`, X_2 = `men_some_college`.

7) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); X_1 = `women_bachelors_degree`, X_2 = `men_bachelors_degree`.

8) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = women_advanced_degree, X_2 = men_advanced_degree.

9) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_women_16-24, X_2 = black_men_16-24.

10) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_women_25-54, X_2 = black_men_25-54.

11) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_women_55-64, X_2 = black_men_55-64.

12) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_women_65+, X_2 = black_men_65+.

13) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_16-24, X_2 = hispanic_16-24.

14) Набір даних used_cars_UA.csv (для автомобілей на дизелі); X= пробіг (у милях), Y=ціна (у фунтах)

15) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_25-54, X_2 = hispanic_25-54.

16) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = black_55-64, X_2 = hispanic_55-64.

17) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = white_women_16-24, X_2 = white_men_16-24.

18) Набір даних employment_to_population_1979_to_2022 (Співвідношення зайнятості до населення США); X_1 = white_women_25-54, X_2 = white_men_25-54.

19) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); $X_1 =$ `white_women_55-64`, $X_2 =$ `white_men_55-64`.

20) Набір даних `employment_to_population_1979_to_2022` (Співвідношення зайнятості до населення США); $X_1 =$ `white_women_65+`, $X_2 =$ `white_men_65+`.

Завдання 2. Побудувати лінійну регресійну залежність між висотою над рівнем моря та відповідною залежною ознакою Y. Побудувати оптимальну нелінійну модель. За різними критеріями дослідити її на точність. Дані містяться у відповідному файлі.

1) $0109.txt$, $Y =$ атмосферний тиск.
2) $0203.txt$, $Y =$ температура.
3) $2510.txt$, $Y =$ вологість.
4) $2705.txt$, $Y =$ температура.
5) $2603.txt$, $Y =$ атмосферний тиск.
6) $0301.txt$, $Y =$ вологість.
7) $0304.txt$, $Y =$ температура.
8) $0308.txt$, $Y =$ атмосферний тиск.
9) $0403.txt$, $Y =$ вологість.
10) $0408.txt$, $Y =$ температура.
11) $0410.txt$, $Y =$ атмосферний тиск.
12) $0501.txt$, $Y =$ вологість.
13) $0503.txt$, $Y =$ температура.
14) $0506.txt$, $Y =$ атмосферний тиск.
15) $0702.txt$, $Y =$ вологість.
16) $0703.txt$, $Y =$ температура.
17) $0705.txt$, $Y =$ атмосферний тиск.

18) 0908.txt, Y = вологість.

19) 0910.txt, Y = температура.

20) 0911.txt, Y = атмосферний тиск.

Завдання 3. За даними із файлу `Результати замірів діаметрів і висот дерев у ялицевому насадженні.xslx` обчислити коефіцієнт кореляції між діаметрами ($\mathrm{x_i}$) і висотами (y_i) 20 дерев, дати оцінку. Побудувати пряму регресії, оцінити значущість коефіцієнтів регресії прямої.

Література

[1] *Біометрія. Методичні рекомендації до виконання лабораторних робіт з дисципліни для студентів денної форми навчання за освітнім ступенем бакалавр зі спеціальностей 205 «Лісове господарство», 201 «Агрономія»*, Укладач : К. М. Кудряшова, Г.І. Рябуха, Л.А. Шевченко. Чернігів НУ «Чернігівська політехніка», 2020. 32с.

[2] *Біометрія: навчальний посібник*/ Укладач: С.С. Чепур. Ужгород: Вид-во УжНУ «Говерла», 2023. 196 с.

[3] Бучавий Ю.В., Рудченко А.Г. *Практикум з біометрії: методичні рекомендації для студентів спеціальностей 091 «Біологія», 101 «Екологія» та 183 «Технології захисту навколишнього середовища»*. НТУ «Дніпровська політехніка». Дніпро: НТУ «ДП», 2019. 43 с.

[4] Вітер Р.М., Шпарик Ю.С. *Методичні вказівки до практичних занять з «Біометрії» для студентів II курсу напряму підготовки 6.090103 «Лісове і садовопаркове господарство»*. Івано-Франківськь: Територія друку, 2013. 38 с.

[5] Зайцев Є. П. *Теорія ймовірностей і математична статистика*. Київ: Алерта, 2013. 440 с.

[6] Калінін М.І. Єлісєєв В.В. *Біометрія: Підручник для студентів вузів біологічних і екологічних напрямків*. Миколаїв: Вид-во МФ НаУКМА, 2000. 204 с.

[7] Чепур С.С. *Біометрія: методичний посібник*. Ужгород: Видавництво УжНУ «Говерла», 2020. 40 с.

[8] *Біометрія: навчальний посібник* / Укладач: С.С. Чепур. Ужгород: Вид-во УжНУ «Говерла», 2023. 196 с.

доц. **Тегза Антоніна Михайлівна** – канд. фіз.-мат. наук;

доц. **Синявська Ольга Олександрівна** – канд. фіз.-мат. наук.

Біометрія: методичні вказівки до виконання типових індивідуальних завдань.