МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ МІСЬКОГО ГОСПОДАРСТВА імені О. М. БЕКЕТОВА

МЕТОДИЧНІ ВКАЗІВКИ

до самостійної роботи
з курсу

ТЕОРІЯ ІМОВІРНОСТЕЙ

(для студентів 2 курсу заочної форми навчання ФПО та ЗН напряму підготовки 6.060101 «Будівництво» спеціальності «Міське будівництво і господарство»)

Харків
ХНУМГ
2014

Методичні вказівки до самостійної роботи з курсу «Теорія імовірностей» (для студентів 2 курсу заочної форми навчання ФПО та ЗН напряму підготовки 6.060101 «Будівництво» спеціальності «Міське будівництво і господарство») / Харк. нац. ун-т міськ. госп-ва ім. О. М. Бекетова; уклад.: В. М. Охріменко, Т. Б. Воронкова, О. О. Воронков. - Х.: ХНУМГ, 2014. - 40 с.

Укладачі : доц. В. М. Охріменко,
ст. викл. Т. Б. Воронкова,
ст. викл. О. О. Воронков.

Рекомендовано кафедрою „Інформаційних систем і технологій в міському господарстві”, протокол № 88 від 11.05.12 р.

ЗМІСТ

Загальні відомості 4
ЗМ 1. ТЕОРІЯ ІМОВІРНОСТЕЙ 5
Тема 1. Основні поняття теорії імовірностей (1 година) 5
Тема 2. Операції над подіями. Теореми теорії імовірностей. Основні формули теорії імовірностей (2 години) 6
Тема 3. Випадкова величина та її закони розподілу (1 година) 8
Тема 4. Числові характеристики випадкової величини (1 година) 9
Тема 5. Найважливіші для практики закони розподілу випадкових величин (3 години) 11
Тема 6. Система випадкових величин. Закони розподілу та числові характеристики системи (1 година) 13
Тема 7. Закон великих чисел (1 година) 18
ЗМ 2. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ 18
Тема 8. Обробка статистичних даних (8 годин) 18
Тема 9. Елементи теорії кореляції (14 годин) 27
Тема 10. Перевірка статистичних гіпотез (14 години) 33
ЗМ 3. ВИПАДКОВІ ПРОЦЕСИ 34
Тема 11. Елементи теорії випадкових процесів (16 годин) 34
Список джерел 36
Додатки 37

ЗАГАЛЬНІ ВІДОМОСТІ

Курс «Теорія імовірностей» є нормативною дисципліною у навчальному плані напряму «Будівництво». Обсяг курсу становить 72 академічних години або 2 кредити ECTS, тому числі обсяг самостійної роботи студента складає 64 академічних години.

Відповідно до програми курс розділений на три змістових модуля: «Теорія імовірностей», «Математична статистика» та «Випадкові процеси». Метою вивчення дисципліни «Теорія імовірностей» є формування базових знань в області застосування імовірнісно-статистичного апарата, вивчення закономірностей у масових випадкових явищах, визначення їх імовірнісних характеристик з метою прогнозування.

В результаті вивчення курсу студенти повинні оволодіти основними методами визначення імовірнісних характеристик випадкових величин, статистичного опису результатів спостереження та перевірки статистичних гіпотез для прийняття на їх основі обґрунтованих рішень. У методичних вказівках до самостійної роботи для кожної теми зазначено обсяг витрат часу на вивчення, що відповідає програмі курсу. Наприкінці методичних вказівок наведено список основних і додаткових підручників, які рекомендується використовувати. Кожна тема супроводжується посиланнями на відповідні їй сторінки підручників. Після вивчення теоретичного матеріалу треба дати відповіді на контрольні запитання з теми, а також вирішити задачі, пропоновані для самостійного розв'язання. Для полегшення роботи перед задачами для самостійного розв'язання наведене розв'язання аналогічних прикладів.

ЗМ 1. ТЕОРІЯ ІМОВІРНОСТЕЙ

Тема 1. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІМОВІРНОСТЕЙ (1 година)

Основні поняття та визначення теорії імовірностей: випадкові й детерміновані явища, дослід, випадкова подія, імовірність. Класичний та статистичний методи визначення імовірності випадкової події. Поняття про частоту випадкової події.

Література: [1] с. 7-16; [3] с. 12-16; [4] с. 15-34.

Запитання для самоперевірки:

1. Дайте визначення випадкової події.
2. Які події називаються: а) достовірними? б) рівноможливими? в) несумісними? г) протилежними? Наведіть приклади.
3. Чи є протилежні події несумісними?
4. Чи є несумісні події протилежними?
5. Дайте визначення імовірності випадкової події.
6. Як підрахувати імовірність події класичним методом?
7. Що розуміють під повною групою подій? Наведіть приклади.
8. Чи завжди можна визначити імовірність випадкової події класичним методом?
9. Як пов'язані між собою імовірність і частота появи події?

Приклад 1.1 Студент прийшов здавати іспит. Він знає 15 з 20 питань програми. Визначити імовірність того, що він відповість на три пропонованих екзаменаційних питання.

Розв'язання

Запишемо подію, імовірність якої необхідно визначити, А={студент знає відповіді на три запитання}. Виразимо її через елементарні події:

A_1 ={знає відповідь на перше запитання}

A_2 ={знає відповідь на друге запитання}

A_3 ={знає відповідь на третє запитання}

$$A = A_1 * A_2 * A_3.$$

Події A_1, A_2, A_3 – залежні події. Обчислимо їх умовні імовірності.

$$P(A_1) = \frac{15}{20}.$$

Умовна імовірність події A_2, за умови, що відбулася подія A_1,

$$P(A_2|A_1)= \frac{14}{19}.$$

Умовна імовірність події A_3, за умови, що відбулися події A_1 і A_2,

$$P(A_3|A_1 * A_2)= \frac{13}{18}.$$

Тоді імовірність події А за теоремою множення:

$$P(A) = P(A_1) * P(A_2|A_1) * P(A_3|A_1 * A_2) = \frac{15}{20} * \frac{14}{19} * \frac{13}{18} = 0{,}368.$$

Задачі для самостійного розв'язання:

1.1. На складі 20 деталей, з яких 17 придатних. Визначити імовірність того, що з трьох навмання взятих деталей всі виявляться придатними.

1.2. На складі 50 придатних і 5 дефектних деталей. Визначити імовірність того, що серед п'яти навмання взятих деталей одна виявиться дефектною.

1.3. Учасники жеребкування тягнуть із ящика жетони з номерами від 1 до 100. Знайти імовірність того, що номер першого жетона, який навмання витягнутий з ящика, не містить цифру 5.

1.4. У партії з 20 готових виробів є 4 бракованих. Партію ділять на дві рівні частини. Визначити імовірність того, що браковані вироби розділяться нарівно.

1.5. Набираючи номер телефону, абонент забув останні дві цифри і, пам'ятаючи, що вони різні, набрав їх навмання. Знайти імовірність того, що були набрані потрібні дві цифри.

1.6. У розіграші першості з баскетболу беруть участь 18 команд, з яких випадковим способом формуються дві групи по 9 команд у кожній. Серед учасників змагань є 5 команд екстракласу. Знайти імовірність того, що а) всі команди екстракласу потраплять в одну групу; б) дві команди потраплять в одну з груп, а три - в іншу.

1.7. Є дві урни, у першій з яких a білих та b чорних кулі, у другій – c білих та d чорних. З кожної урни виймають по одній кулі. Знайти імовірність того, що обидві вийнятих кулі опиняться білими.

1.8. У ліфт будинку, в якому сім поверхів, на першому поверсі ввійшли три пасажири. Кожний з них з однаковою імовірністю виходить на кожному з поверхів. Знайти імовірність того, що всі пасажири вийдуть одночасно (на тому самому поверсу).

Тема 2. ОПЕРАЦІЇ НАД ПОДІЯМИ. ТЕОРЕМИ ТЕОРІЇ ІМОВІРНОСТЕЙ. ОСНОВНІ ФОРМУЛИ ТЕОРІЇ ІМОВІРНОСТЕЙ (2 години)

Сума та добуток випадкових подій. Теорема додавання для несумісних та сумісних подій. Протилежні події. Теорема множення для залежних та незалежних подій. Залежні події. Умовна імовірність випадкової події. Формула повної імовірності. Формула Бейєса (теорема гіпотез). Повторні незалежні випробування. Локальна та інтегральна теореми Лапласа. Формула Пуассона.

Література: [1] С. 17-20; 23-28; 31-51; [3] С. 17-33; [4] С. 37-81.

Запитання для самоперевірки:

1. Як визначити імовірність суми сумісних подій?
2. Чи може сума двох подій збігатися з їхнім добутком?
3. Наведіть приклади залежних і незалежних подій.
4. Що розуміють під умовною імовірністю події?
5. Як визначається імовірність добутку двох подій?
6. В яких випадках для визначення імовірності застосовується формула Бернуллі?

7. Дайте визначення найімовірнішого числа появ події.

8. Як обчислити найімовірніше число появ події?

9. Чим розрізняються задачі, в яких потрібне застосування локальної і інтегральної граничних теорем?

10. В яких випадках замість формули Бернуллі використовується формула Пуассона?

Приклад 2.1. Є три однакові на вигляд урни. У першій – 2 білі та 3 чорних кулі, в другій – 4 білі та 1 чорна, у третій – 3 білі. Навмання з однієї з урн виймають одну кулю. Визначити імовірність того, що вийнята куля виявиться білою.

Розв'язання

Позначимо подію А={вийнята куля біла}. Висуваємо три гіпотези:

H_1 ={обрана перша урна}

H_2 ={обрана друга урна}

H_3 ={обрана третя урна}

Імовірність кожної гіпотези

$$P(H_1) = P(H_2) = P(H_3) = 1/3.$$

Умовні імовірності події А:

$$P(A|H_1)= \frac{2}{5}; \qquad P(A|H_2)= \frac{4}{5}; \qquad P(A|H_3)= 1.$$

Повна безумовна імовірність події А

$$P(A)= P(H_1)*P(A|H_1)+P(H_2)*P(A|H_2)+P(H_3)*P(A|H_3) =$$

$$=\frac{1}{3}*\frac{2}{5} + \frac{1}{3}*\frac{4}{5} + \frac{1}{3}*1=\frac{11}{15}.$$

Задачі для самостійного розв'язання:

2.1. На кожній з 5 однакових карток надруковано одну з наступних букв: «О», «П», «Р», «С», «Т». Картки ретельно перемішані. Знайти імовірність того, що на вийнятих по одній та розташованих в одну лінію картках з'явиться слово «СПОРТ».

2.2. В урні *a* білих та b чорних куль. З урни одну за одною виймають дві кулі. Знайти імовірність того, що обидві кулі будуть білі.

2.3. В урні *a* білих та b чорних куль. З урни одну за одною виймають всі кулі, що знаходяться у ній. Знайти імовірність того, що другою буде вийнято білу кулю.

2.4. Два стрілки незалежно один від одного роблять два постріли (кожний у свою мішень). Імовірність влучення у мішень при одному пострілі для першого стрільця p_1, для другого p_2. Вигравшим змагання вважається той стрілець, у мішені якого буде більше пробоїн. Знайти імовірність того, що виграє перший стрілець.

2.5. Імовірність того, що протягом однієї зміни виникне неполадка верстата, дорівнює 0,05. Визначити імовірність того, що протягом трьох змін верстат не зламається жодного разу.

2.6. Прилади одного найменування виготовляють два заводи. На виробництво, що використовує прилади, 2/3 приладів надходять з першого заводу

та 1/3 – з другого. Імовірність безвідмовної роботи приладу (надійність), виготовленого першим заводом, дорівнює 0,95, а другим – 0,9. Визначити повну надійність приладу, що надійшов на виробництво.

2.7. Службовці підприємства розподілені за підрозділами та статтю в такий спосіб: у виробничому відділі – 8 жінок і 18 чоловіків, у плановому відділі - 4 жінки і 9 чоловіків, у відділі реалізації – 10 жінок і 5 чоловіків. Навмання обраний службовець виявився чоловіком. Визначити імовірність того, що він працівник а) виробничого відділу; б) планового відділу; в) відділу реалізації.

2.8. Готові вироби містять 5% браку. Визначити імовірність того, що в числі п'яти взятих навмання виробів: а) немає жодного дефектного; б) два дефектних.

2.9. Виробництво дає 1% браку. Визначити імовірність того, що з 1500 виробів бракованих буде не більше 20.

Тема 3. ВИПАДКОВА ВЕЛИЧИНА ТА ЇЇ ЗАКОНИ РОЗПОДІЛУ (1 година)

Поняття випадкової величини. Безперервні й дискретні випадкові величини. Закон розподілу випадкової величини. Ряд розподілу. Функція розподілу та її властивості. Щільність розподілу та її властивості.

Література: [1] С. 57-59; 105-117; [3] С. 34-39; [4] С. 82-99.

Запитання для самоперевірки:

1. Дайте визначення випадкової величини.
2. Яка випадкова величина називається дискретною? Наведіть приклади.
3. Яка випадкова величина називається безперервною? Наведіть приклади.
4. Поясніть, з якою метою в теорії імовірностей розрізняють дискретні і безперервні випадкові величини?
5. Що має на увазі термін «закон розподілу»? В яких формах може бути представлений закон розподілу випадкової величини?
6. Чи може функція розподілу бути: а) більше одиниці; б) від’ємною?
7. Що розуміють під щільністю розподілу випадкової величини?
8. Чому не має сенсу поняття щільності розподілу для дискретної випадкової величини?
9. Яка розмірність щільності розподілу?
10. Перелічіть властивості щільності розподілу.
11. Як, маючи ряд розподілу, знайти значення функції розподілу?
12. Як виражається імовірність влучення випадкової величини на інтервал значень, якщо відомо функцію розподілу? Щільність розподілу?

Приклад 3.1. Випадкову величину X задано функцією розподілу:

$$F(x) = \begin{cases} 0, x \le 0 \\ x^2 / 4, 0 < x \le 2 \\ 1, x > 2 \end{cases}.$$

Знайти щільність розподілу імовірностей та побудувати графіки функцій f(x) та F(x).

Розв'язання

Знайдемо щільність імовірностей, взявши похідну від функції розподілу на кожному інтервалі:

$$f(x)=\begin{cases}0, x\le 0\\ x/2, 0<x\le 2\\ 0, x>2\end{cases}.$$

в) Побудуємо графіки функцій f(x) та F(x).

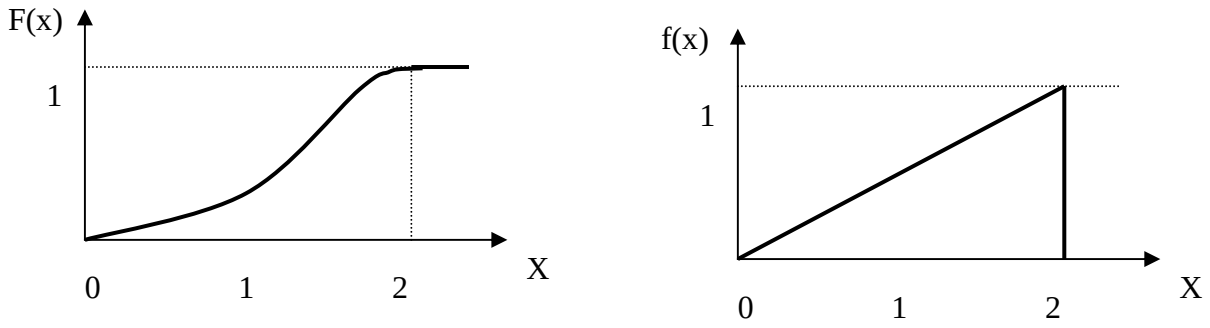

Рис. 3.1 - Графік функції розподілу та щільності імовірностей

Задачі для самостійного розв'язання:

3.1. Розглядаючи невипадкову величину С як окремий вид випадкової, побудувати для неї функцію розподілу, знайти математичне сподівання і дисперсію.

3.2. У партії з 30 виробів є 7 дефектних. З цієї партії випадковим способом обрані три вироби для перевірки їх якості. Побудувати ряд розподілу кількості відібраних для перевірки виробів (випадкової величини X).

3.3. Для умов попередньої задачі побудувати функцію розподілу кількості відібраних для перевірки виробів (випадкової величини X).

3.4. Випадкову величину X задано функцією розподілу

$$F(x)=\begin{cases}0, & x<2\\ (x-2)^2, & 2\le x\le 3\\ 1, & x>3\end{cases}.$$

Знайти щільність розподілу f(x) та визначити імовірність влучення випадкової величини X в інтервал (1,5; 2).

Тема 4. ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВОЇ ВЕЛИЧИНИ (1 година)

Середнє значення та математичне сподівання. Мода. Медіана. Моменти випадкової величини: початкові й центральні. Дисперсія та середнє квадратичне відхилення. Теореми про числові характеристики.

Література: [1] С. 67-70; 77-84; [3] С. 40-45; [4] С. 107-121.

Запитання для самоперевірки:

1. Назвіть основні числові характеристики випадкових величин.

2. Як пов'язані між собою математичне сподівання і середнє арифметичне значень випадкової величини?

3. Математичне сподівання - випадкова величина чи ні?

4. Чи є дисперсія випадковою величиною?

5. Як математичне сподівання і дисперсія характеризують випадкову величину?

6. Чим зручне застосування замість дисперсії середнього квадратичного відхилення?

7. В яких одиницях вимірюють математичне сподівання?

8. В яких одиницях вимірюють дисперсію?

9. Чому дорівнює математичне сподівання невипадкової величини С?

10. Як мода і медіана характеризують випадкову величину?

Приклад 4.1. Нехай випадкову величину Х задано функцією розподілу:

$$F(x) = \begin{cases} 0, x \le 0 \\ x^2/4, 0 < x \le 2 \\ 1, x > 2 \end{cases}.$$

Визначити числові характеристики безперервної випадкової величини.

Розв'язання

Знайдемо щільність імовірностей:

$$f(x) = \begin{cases} 0, x \le 0 \\ x/2, 0 < x \le 2 \\ 0, x > 2 \end{cases}.$$

Визначимо математичне сподівання Х:

$$M[X] = \int_0^2 x * x/2dx = 1/2\int_0^2 x^2 dx = \frac{1}{2}\frac{x^3}{3}\Big|_0^2 = 4/3.$$

Знайдемо дисперсію Х:

$$D[X] = \int_0^2 x^2 * x/2dx - (4/3)^2 = 1/2\int_0^2 x^3 dx - (4/3)^2 = \frac{1}{2}\frac{x^4}{4}\Big|_0^2 - (4/3)^2 = 2/9.$$

Задачі для самостійного розв'язання:

4.1. Для умов прикладу 3.2 визначити математичне сподівання, дисперсію і середнє квадратичне відхилення випадкової величини.

4.2. Для умов прикладу 3.4 визначити математичне сподівання, дисперсію і середнє квадратичне відхилення випадкової величини.

4.3. До випадкової величини Х додали невипадкову величину *a*. Як від цього зміняться її числові характеристики: математичне сподівання, дисперсія і середнє квадратичне відхилення?

4.4. Випадкову величину Х помножили на невипадкову величину *a*. Як від цього зміняться її числові характеристики: математичне сподівання, дисперсія і середнє квадратичне відхилення?

4.5. Виконують два незалежних постріли по мішені. Імовірність влучення при кожному пострілі дорівнює р. Розглядаються дві випадкові величини: Х- різниця між числом влучень і числом промахів і Y - сума числа влучень і числа промахів. Побудувати для випадкових величин Х і Y ряд розподілу (для кожної окремо) і знайти їхні числові характеристики.

Тема 5. НАЙВАЖЛИВІШИ ДЛЯ ПРАКТИКИ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН (3 години)

Біноміальний закон розподілу. Закон розподілу Пуассона. Експонентний закон розподілу. Поняття найпростішого потоку подій. Число подій, що потрапляють на ділянку часу τ. Проміжок часу між двома сусідніми подіями в найпростішому потоці Т. Нормальний закон розподілу імовірностей. Інтеграл імовірностей. Правило трьох сигма. Поняття про центральну граничну теорему. Закон рівномірної щільності.

Література: [1] С. 60-67; 118-131; 146-150; [3] С. 46-61; [4] С. 129-150; 153-167.

Запитання для самоперевірки:

1. Яким умовам повинні задовольняти повторні незалежні випробування?
2. Як визначають числові характеристики випадкової величини, розподіленої за законом Бернуллі?
3. Який зв'язок існує між біноміальним і пуассонівським розподілами?
4. Яким умовам повинна задовольняти випадкова величина, підпорядкована закону Пуассона?
5. Як визначають числові характеристики закону розподілу Пуассона?
6. Якими параметрами визначається експонентний закон розподілу випадкової величини?
7. Чому дорівнює щільність імовірності випадкової величини з нормальним законом розподілу?
8. Якими параметрами визначається нормальний закон розподілу випадкової величини?
9. Як змінюється графік нормального закону із зміною середнього квадратичного відхилення випадкової величини?
10. Як визначити імовірність влучення нормально розподіленої випадкової величини на задану ділянку?
11. Поясніть імовірнісний смисл параметрів нормального розподілу.
12. Поясніть смисл центральної граничної теореми.

Приклад 5.1. Космічні частки, що потрапляють у супутник, утворюють поле із щільністю λ=1 частка/м^2. Агрегат супутника, який знаходиться у полі часток, займає площу s=10 см^2. Для виходу з ладу агрегату свідомо досить влучення в нього двох часток. При влученні однієї частки він виходить з ладу з імовірністю p=0,5. Визначити імовірність виходу з ладу агрегату.

Розв'язання

Позначимо подію, що цікавить нас, А={вихід агрегату з ладу}. Цій події відповідають дві гіпотези:

H_1 ={в агрегат потрапила одна частка},

H_2 ={в агрегат потрапило дві частки}.

Умовні імовірності події А: $P(A/H_1)=0,5$, $P(A/H_2)=1$.

Імовірності гіпотез визначимо за законом розподілу Пуассона, параметр

якого a=1*0,001=0,001:

$$P(H_1) = P(1) = \frac{0{,}001^1}{1!} e^{-0{,}001} = 0{,}000999 \quad ,$$

$$P(H_2) = 1 - P(0) - P(1) = 1 - \frac{0{,}001^0}{0!} e^{-0{,}001} - P(1) = 1 - 0{,}999 - 0{,}000999 = 10^{-6} .$$

За формулою повної ймовірності дістанемо ймовірність події А:

$P(A)=P(H_1)P(A/H_1)+ P(H_2)P(A/H_2) = 0{,}000999*0{,}5+10^{-6}*1=5{,}005*10^{-4}$.

Приклад 5.2. У нормально розподіленій сукупності 15% значень X менше за 12 та 40% значень X більше за 16,2. Знайти середнє значення та середнє квадратичне відхилення даного розподілу.

Розв'язання

З умови задачі витікає, що

P{X<12}=0,15 та P{X<16,2}=0,6.

Запишемо систему рівнянь:

$$\begin{cases} \Phi\left(\frac{12-m_x}{\sigma_x}\right)+0{,}5=0{,}15 \\ \Phi\left(\frac{16{,}2-m_x}{\sigma_x}\right)+0{,}5=0{,}6 \end{cases} \Rightarrow \begin{cases} \Phi\left(\frac{12-m_x}{\sigma_x}\right)=-0{,}35 \\ \Phi\left(\frac{16{,}2-m_x}{\sigma_x}\right)=0{,}1 \end{cases} \Rightarrow \begin{cases} \frac{12-m_x}{\sigma_x}=-1{,}04 \\ \frac{16{,}2-m_x}{\sigma_x}=0{,}25 \end{cases} \Rightarrow . \Rightarrow \begin{cases} m_x=15{,}386 \\ \sigma_x=3{,}256 \end{cases} .$$

Параметри розподілу: m_x = 15,386, σ_x =3,256.

Задачі для самостійного розв'язання:

5.1. Випадкова величина X підлегла закону розподілу Пуассона з математичним сподіванням a=3. Побудувати многокутник розподілу і функцію розподілу випадкової величини X. Знайти імовірність того, що випадкова величина X прийме значення менше, ніж її математичне сподівання.

5.2. Випадкова величина X підпорядкована експонентному закону розподілу з параметром μ:

$$f(x) = \begin{cases} \mu e^{-\mu x} & при\ x > 0 \\ 0 & при\ x < 0 \end{cases}$$

Побудувати криву розподілу, визначити функцію розподілу і знайти імовірність того, що випадкова величина X прийме значення, менше її математичного сподівання.

5.3. Вимірювальний прибор має систематичну помилку 5 м і середню квадратичну помилку 75 см. Яка імовірність того, що помилка вимірювання не перевершить за абсолютною величиною 5 м?

5.4. Випадкова величина X підпорядкована нормальному закону з математичним сподіванням, рівним нулю. Імовірність влучення цієї випадкової величини на ділянку від $-\alpha$ до $+\alpha$ дорівнює 0,5. Знайти середнє квадратичне відхилення і написати вираз нормального закону.

5.5. У світлофорі на перехресті 1 хвилину горить зелене світло та 0,5 хвилини – червоне. Автомобіль під'їжджає до перехрестя у випадковий момент, не пов'язаний з роботою світлофора. Знайти імовірність того, що він проїде перехрестя, не зупиняючись.

Тема 6. СИСТЕМА ВИПАДКОВИХ ВЕЛИЧИН. ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ СИСТЕМИ (1 година)

Багатовимірна випадкова величина. Поняття системи випадкових величин. Система двох випадкових величин. Функція розподілу та щільність розподілу імовірностей системи двох випадкових величин, їх властивості. Числові характеристики системи, кореляційний момент та коефіцієнт кореляції. Функції випадкових величин.

Література: [1] С. 135-142; 153-164; 177-181; [3] С. 62-72; [4] С. 177-190; 213-219; 258-276.

Запитання для самоперевірки:

1. Що являє собою багатомірна випадкова величина?
2. Що являє собою функція розподілу системи двох випадкових величин? Перелічіть її властивості.
3. Перелічіть числові характеристики системи двох випадкових величин.
4. Що характеризує кореляційний момент системи двох випадкових величин?
5. Для чого використовується коефіцієнт кореляції?
6. Перелічіть теореми про числові характеристики.
7. Чому дорівнює середнє квадратичне відхилення добутку невипадкової величини С на випадкову величину Х?
8. Сформулюйте теорему додавання математичних сподівань для випадкових величин: а) залежних і незалежних; б) корельованих і некорельованих.
9. Чому дорівнює математичне сподівання добутку двох незалежних випадкових величин?

Приклад 6.1. Матриця розподілу випадкового вектора (X, Y) має вигляд:

y_i / x_i	0	2	5
1	0,1	0	0,2
2	0	0,3	0
4	0,1	0,3	0

Знайти числові характеристики.

Розв'язання

Для визначення числових характеристик випадкової величини Х побудуємо її ряд розподілу.

x_i	1	2	4
p_i	0,3	0,3	0,4

де P{X=1}=0,1+0+0,2=0,3; P{X=2}=0+0,3+0=0,3; P{X=4}=0,1+0,3+0=0,4.

Тоді $m_x = \Sigma x_i p_i = 1*0,3+2*0,3+4*0,4=2,5$;

$D_x = \Sigma(x_i - m_x)^2 p_i = (1-2,5)^2*0,3+(2-2,5)^2*0,3+(4-2,5)^2*0,4=1,65$;

$\sigma_x = \sqrt{D_x} = \sqrt{1,65} = 1,285$.

Аналогічно побудуємо ряд розподілу випадкової величини Y.

y_i	0	2	5
p_i	0,2	0,6	0,2

та визначимо її числові характеристики:

$m_y = \Sigma y_i p_i = 0*0,2+2*0,6+5*0,2=2,2;$

$D_y = \Sigma(y_i - m_y)^2 p_i = (0-2,2)^2*0,3+(2-2,2)^2*0,6+(5-2,2)^2*0,2=2,54;$

$\sigma_y = \sqrt{D_y} = \sqrt{2,54} = 1,6.$

Визначимо кореляційний момент системи випадкових величин X, Y:

$$K_{xy} = \sum_{i=1}^{3}\sum_{j=1}^{3}(x_i - m_x)(y_j - m_y)p_{ij} = (1-2,5)[(0-2,2)*0,1+(2-2,2)*0+(5-2,2)*0,2]+$$
$$+(2-2,5)[(0-2,2)*0+(2-2,2)*0,3+(5-2,2)*0]+$$
$$+(4-2,5)[(0-2,2)*0,1+(2-2,2)*0,3+(5-2,2)*0]=-0,9.$$

Визначимо коефіцієнт кореляції:

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{-0,9}{1,285*1,6} = -0,438.$$

Приклад 6.2 Двічі кидають гральну кістку. Випадкові величини X - число появ шістки, Y - число появ парної цифри. Описати закони розподілу випадкових величин X та Y; описати закон розподілу випадкового вектора (X, Y); встановити, чи є залежними X та Y.

Розв'язання

Знайдемо закон розподілу випадкового вектора (X, Y), склавши матрицю розподілу, оскільки X та Y – дискретні.

y_i / x_i	0	1	2	$P\{X=x_i\}$
0	1/4	1/3	1/9	25/36
1	0	1/6	1/9	10/36
2	0	0	1/36	1/36
$P\{Y=y_i\}$	1/4	1/2	1/4	1

Для визначення умовних імовірностей $P_{ij} = P\{X=x_i/Y=y_j\}$ скористуємося теоремами додавання та множення імовірностей.

Дістанемо ряд розподілу випадкової величини X, для чого підсумовуємо умовні імовірності у кореляційній таблиці за рядками, та Y, для чого підсумовуємо умовні імовірності за стовпцями.

Встановимо, чи залежні X та Y. Відомо, що у випадку незалежних подій умовна імовірність події дорівнює її безумовній імовірності, але рівність $P\{y_j /x_i\} = P\{y_j\}$ не дотримується. Наприклад,

$$P_{22}\{y_j=1, x_i=1\}= P\{x_i=1\}*P\{y_j=1/x_i=1\}=1/6.$$

Отже, X та Y залежні.

Приклад 6.3 Для умови попередньої задачі визначити числові характеристики випадкового вектора (X, Y).

Розв'язання

Визначимо математичні сподівання X та Y:

$$m_x = \Sigma x_i p_i = 0*\frac{25}{36}+1*\frac{5}{18}+2*\frac{1}{36}=\frac{1}{3};$$

$$m_y = \Sigma y_i p_i = 0*\frac{1}{4}+1*\frac{1}{2}+2*\frac{1}{4}=1.$$

Визначимо дисперсії X та Y:

$$D_x = \Sigma(x_i - m_x)^2 p_i = \frac{5}{18};$$

$$D_y = \Sigma(y_i - m_y)^2 p_i = \frac{1}{2}.$$

Визначимо кореляційний момент системи випадкових величин X, Y:

$$K_{xy} = \sum_{i=1}^{3}\sum_{j=1}^{3}(x_i - m_x)(y_j - m_y)p_{ij} = (0-\frac{1}{3})\left[(0-1)*\frac{1}{4}+(1-1)*\frac{1}{3}+(2-1)*\frac{1}{9}\right]+$$

$$+(1-\frac{1}{3})\left[0+(1-1)*\frac{1}{6}+(2-1)*\frac{1}{9}\right]+$$

$$+(2-\frac{1}{3})\left[(0-1)*0+(1-1)*0+(2-1)*\frac{1}{36}\right]=\frac{1}{6}.$$

Визначимо коефіцієнт кореляції:

$$r_{xy} = \frac{K_{xy}}{\sqrt{D_x D_y}} = \frac{\frac{1}{6}}{\sqrt{\frac{5}{18}*\frac{1}{6}}} = \frac{1}{\sqrt{5}}.$$

Приклад 6.4 На вимірювальний прибор надходить випадковий вектор (X, Y) з наступними характеристиками: m_x=-1; m_y=1; σ_x=2; σ_y=3; r_{xy}=0,5. На виході прибору вимірюють величину $Z=(X-Y)^2$. Визначити математичне сподівання випадкової величини Z.

Розв'язання

Оскільки коефіцієнт кореляції не дорівнює нулю, випадкові величини X та Y корельовані. Перетворимо вираз Z:

$$M[Z] = M[(X-Y)^2] = M[X^2-2XY+Y^2]=$$

математичне сподівання суми випадкових величин дорівнює сумі математичних сподівань:

$$= M[X^2]- M[2XY]+ M[Y^2]=$$

другий початковий момент запишемо як суму дисперсії та квадрата математичного сподівання і врахуємо, що математичне сподівання добутку корельованих випадкових величин дорівнює добутку їх математичних сподівань плюс кореляційний момент:

$$= D_x+M^2[X] - 2(M[X]*M[Y]+K_{xy})+ D_y+M^2[Y] =$$

$$= 2^2+(-1)^2 - 2*(-1*1+2*3*0,5) + 3^2+1^2 = 11.$$

Приклад 6.5 Визначити математичне сподівання та дисперсію випадкової величини Z, якщо Z=3+4(X-Y). Числові характеристики: m_x=-2; m_y=4; D_x=4; D_y=9; r_{xy}=-0,5.

Розв'язання

Оскільки коефіцієнт кореляції не дорівнює нулю, випадкові величини X та Y корельовані. Визначимо математичне сподівання Z:

$$M[Z] = M[3+4(X-Y)] = M[3]+4M[X+Y]=$$
$$= 3+4m_x+4m_y=3+4*(-2)+4*4=11.$$

Визначимо дисперсію Z:

$$D[Z]= D[3+4(X-Y)] = D[3]+16D[X+Y]=$$
$$= 0+16(D[X]+D[Y]+2K_{xy}) =$$
$$=16*[4+9+2*(-0,5)(4*9] = 112,$$

де $K_{xy}=r_{xy}*\sqrt{D_xD_y}$.

Середнє квадратичне відхилення $\sigma_z = \sqrt{D_z} = \sqrt{112} = 10,6$.

Приклад 6.6 Випадкові величини X та Y незалежні та мають наступні характеристики: $m_x=1$; $m_y=2$; $\sigma_x=1$; $\sigma_y=2$. Обчислити математичне сподівання випадкових величин:

а) $U=X^2+2Y^2-XY-4X+Y+4$;

б) $V=(X+Y-1)^2$.

Розв'язання

а) Визначимо математичне сподівання випадкової величини U:

$$M[U]=M[X^2]+2M[Y^2]-M[XY]-4M[X]+M[Y]+M[4] =$$
$$= D[X]+M^2[X]+2(D[Y]+M_2[Y])-M[X]*M[Y]-4M[X]+M[Y]+4=$$
$$1+1+2*(4+4)-1*2-4*1+2+4 = 18;$$

б) Визначимо математичне сподівання випадкової величини V:

$$M[V] = M[(X+Y-1)^2]= M[X^2+2Y-2+Y^2-2Y+1] =$$
$$= M[X^2]+2M[Y]-M[2]+M[Y^2]-2M[Y]+M[1] =$$
$$= D[X]+M^2[X]+2M[Y]-2+D[Y]+M^2[Y]-2M[Y]+1=$$
$$=1+1+4-2+4+4-4+1 = 9.$$

Приклад 6.7. Є випадкова величина X з математичним сподіванням m_x та дисперсією D_x. Знайти математичне сподівання та дисперсію наступних випадкових величин:

а) $Y=-X$;

б) $Z=X+2Y-1$;

в) $U=3X-Y+2Z-3$.

Розв'язання

а) визначимо математичне сподівання випадкової величини Y:

а) $M[Y]=M[-X]= -m_x$;

визначимо дисперсію випадкової величини Y

$$D[Y] = D[-X] = (-1)^2D_x = D_x;$$

б) визначимо математичне сподівання випадкової величини Z:

$$M[Z]=M[X+2Y-1] = m_x -2m_x -1 = - m_x -1;$$

визначимо дисперсію випадкової величини Z з урахуванням того, що математичне сподівання добутку корельованих випадкових величин дорівнює

добутку їх математичних сподівань плюс кореляційний момент:

$$D[Z]=D[X+2Y-1] = D_x + 4D_y + 2K_{xy}.$$

Кореляційний момент запишемо як різницю другого початкового моменту і добутку математичних сподівань, дістанемо

$$K_{xy} = M[X2Y] - M[X] * M[2Y]= M[X(-2X)] - M[X] * M[2Y] =$$
$$= -2M[X^2] + 2M^2[X] = -2(M[X^2] - M^2[X]) = -2D_x.$$

тоді дисперсія Z дорівнюватиме

$$D[Z] = D_x + 4D_x – 4D[X] = D[X].$$

в) Визначимо математичне сподівання випадкової величини Z:

$$M[U]= M[3X-Y+2Z-3] = 3m_x – m_y +2m_z -3 =$$
$$=3m_x + m_x +2(- m_x -1) -3 = 2m_x -5;$$

визначимо дисперсію випадкової величини U

$$D[U]= D[3X-Y+2Z-3] = D[3X] + D[-Y] + D[2Z] + 2(K_{xy}+K_{yz} + K_{xz});$$

визначимо кореляційні моменти

$$K_{xy} = M[3X(-Y)] - M[3X] * M[-Y]= 3M[X^2] - 3M[X] * M[X]=$$
$$= 3M[X^2] - 3M^2[X] = 3D_x;$$
$$K_{xz} = M[3X2Z] - M[3X] * M[2Z]= M[3X2(X+2Y-1)] - M[3X] * M[2X+4Y-2]=$$
$$= M[6X^2 -12X^2-6X)] - 3M[X] *(2M[X]+4M[-X]-2) =$$
$$=M[6X^2] - 12M[X^2]- 6M[X] - 6M^2[X] +12M^2[X] + 6 M[X] =$$
$$=-6 M[X^2] + 6 M^2[X] = -6 (M[X^2] - M^2[X]) = -6D_x.$$
$$K_{yz} = M[-Y2Z] - M[-Y] * M[2Z]= M[2X (X+2Y-1)] - M[X] * M[2(X+2Y-1)]=$$
$$= M[2X^2 -4X^2-2X] - M[X] * (2M[X]-4M[X]-2)=$$
$$=2M[X^2] -4M[X^2]-2M[X] - 2M^2[X] +4M^2[X]+2M[X]=$$
$$=-2M[X^2] + 2M^2[X]=-2(M[X^2] - M^2[X]) = -2D_x.$$

Визначимо тепер дисперсію випадкової величини U:

$$D[U]= D[3X-Y+2Z-3] = 9D[X] + D[X] + 4D[X] + 2(3D_x-6D_x-2D_x)=$$
$$=9D_x + D_x + 4D_x + 6D_x-12D_x-4D_x= 4 D_x.$$

Задачі для самостійного розв'язання:

6.1. Два стрілки незалежно один від іншого проводять по одному пострілу, кожний по своїй мішені. Випадкова величина X – число влучень першого стрільця, Y– число влучень другого стрільця. Імовірність влучення в мішень для першого стрільця дорівнює p_1=0,9, а для другого p_2=0,8. Побудувати функцію розподілу системи випадкових величин X і Y.

6.2. Для умов попереднього прикладу визначити числові характеристики випадкового вектора (X, Y).

6.3. Незалежні випадкові величини X і Y розподілені за нормальними законами з параметрами m_x=2; m_y=-3; σ_x=1; σ_y=2. Визначити імовірність події A={X< m_x і Y<m_y}.

6.4. Відомі математичне сподівання і дисперсія випадкової величини X: m_x=2; D_x=3. Знайти математичне сподівання і дисперсію випадкової величини Y=3X-2.

6.5. Випадкові величини X і Y мають математичні сподівання m_x=-1, m_y=1 і дисперсії D_x=4 і D_y=9. Знайти математичне сподівання випадкової величини Z=3XY + 5.

Тема 7. ЗАКОН ВЕЛИКИХ ЧИСЕЛ (1 година)

Принцип практичної впевненості. Формулювання закону великих чисел. Рівень значущості. Лема Маркова. Нерівність Чебишева. Теорема Чебишева. Теорема Бернуллі.

Література: [1] с. 94-104; [3] с. 73-79; [4] с. 399-413; [5] с. 180-196.

Запитання для самоперевірки:

1. Що називається законом великих чисел? Поясніть смисл цієї назви.
2. Яка роль закону великих чисел у теорії імовірностей?
3. У чому полягає принцип практичної впевненості?
4. Поясніть смисл поняття «рівень значущості».
5. Сформулюйте теорему Чебишева і поясніть, в чому полягає її практичний зміст.
6. Сформулюйте теорему Бернуллі і поясніть, в чому полягає її практичний зміст.
7. Чи можна стверджувати, що при нескінченно великій кількості дослідів n частота події p* дорівнює імовірності цієї події p? Обґрунтуйте відповідь.

Задачі для самостійного розв'язання:

7.1. Вага виробу, що виготовляється підприємством, є випадковою величиною з математичним сподіванням 90 г і дисперсією 0,0225. Визначити імовірність того, що відхилення ваги виробу від її середнього значення за абсолютною величиною не перевищить 0,4 г. Для розв'язання використати нерівність Чебишева.

7.2. З 1000 виробів, що надходять у складальний цех, випадковим способом вибрали 200 виробів для контролю. Серед них виявилося 25 бракованих. Прийнявши частку бракованих виробів з контрольної партії як імовірність виготовлення бракованого виробу, оцінити імовірність того, що у всій партії виробів, які надійшли у складальний цех, бракованих виявиться не менше 10 % і не більше 15 %. Для розв'язання використати теорему Бернуллі.

ЗМ 2. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Тема 8. ОБРОБКА СТАТИСТИЧНИХ ДАНИХ (8 годин)

Вибірковий метод. Поняття генеральної та вибіркової сукупностей. Варіанта. Частота. Варіаційний ряд. Полігон розподілу. Кумулятивна крива. Гістограма розподілу. Визначення закону розподілу спостережуваної ознаки за статистичними даними. Числові характеристики варіаційного ряду. Варіаційний розмах R. Середнє лінійне відхилення d. Дисперсія варіаційного ряду. Стандартне відхилення. Коефіцієнт варіації V. Властивості вибіркових числових характеристик. Довірчий інтервал та довірча імовірність.

Література: [1] С. 185-210; 214-227; [3] С. 80-107; [4] С. 430-444; 451-467; [5] С. 205-207; 217-223; 237-241; 311-316;.

Запитання для самоперевірки:

1. Які завдання вирішує математична статистика? Назвіть основні з них.
2. Поясніть зміст вибіркового методу.

3. У чому полягає різниця між генеральною сукупністю і вибіркою?
4. Яку інформацію про досліджувану ознаку дістають з варіаційного ряду?
5. Що таке оцінка параметра розподілу?
6. Якими властивостями повинні володіти вибіркові числові характеристики варіаційного ряду?
7. Поясніть властивості спроможності й незмещеності оцінок.
8. Чим відрізняються точкова і інтервальна оцінки параметрів розподілу?
9. Поясніть поняття «довірчий інтервал» і «довірча імовірність».

Приклад 8.1. За результатами розпродажу товарів 26 продавцями побудований варіаційний ряд:

Кількість продажів (x_i)	9	12	13	14	15	16	17	19	21	23	27
Кількість продав-ців(m_i)	1	2	3	6	5	3	2	1	1	1	1

Знайти статистичну функцію розподілу F*(x).

Розв'язання

Очевидно, що шукана статистична функція розподілу для всіх x<9 дорівнює нулю (тому що кількість продажів не приймала значень менших за дев'ять)

F*(9)=0.

Визначимо частоту кількості продажів x<12, до них належить кількість продажів 9, причому m(9)=1, тоді

$$F^*(12) = \frac{m(9)}{n} = \frac{1}{26} \approx 0{,}04 .$$

Визначимо частоту кількості продажів x<13, до них належать кількості продажів 9 та 12, причому m(9)=1, а m(12)=2, тоді

$$F^*(13) = \frac{m(9)}{n} + \frac{m(12)}{n} = \frac{1}{26} + \frac{2}{26} \approx 0{,}04 + 0{,}08 = 0{,}12 .$$

Аналогічно отримаємо інші значення статистичної функції розподілу та подамо їх у табличному вигляді.

x_i	9	12	13	14	15	16	17	19	21	23	27	x>27
$F(x_i)$	0	0,04	0,12	0,23	0,46	0,65	0,77	0,85	0,88	0,92	0,96	1

Графік статистичної функції розподілу представлений на рис. 5.1.

Приклад 8.2. З метою дослідження точності приладу зробили 500 вимірювань помилки. Результати вимірювань звели у групований варіаційний ряд:

групи	-4 ...-3	-3 ...-2	-2 ...-1	-1 ...0	0 ...1	1 ...2	2 ...3	3 ...4
m_i	6	25	72	133	120	88	46	10

Побудувати гістограму розподілу та визначити параметри розподілу.

Розв'язання

Визначимо частоти для кожного розряду групованого варіаційного ряду, користуючись формулою

$$p_i^* = \frac{m_i}{n},$$

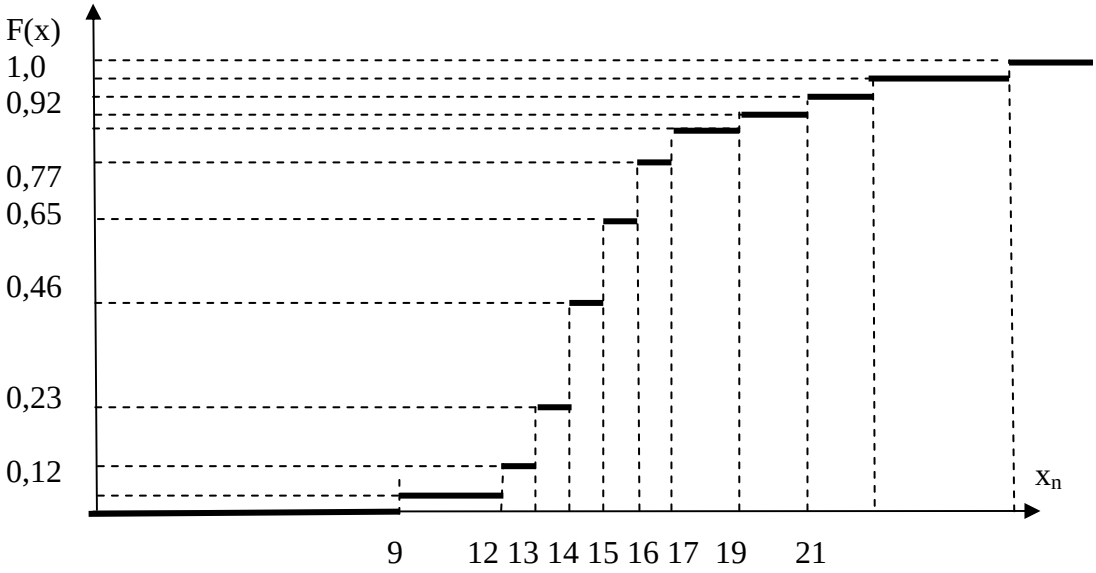

Рис. 8.1 - Статистична функція розподілу.

де m_i – кількість значень помилки X, що потрапили в і-у групу; n – число зроблених вимірів, n = 500.

групи	-4 …-3	-3 …-2	-2 …-1	-1 …0	0 …1	1 …2	2 …3	3 …4
m_i	6	25	72	133	120	88	46	10
p_i*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02

Для побудови гістограми визначимо щільності частот для кожної групи значень за формулою

$$f_i^* = \frac{p_i^*}{l},$$

де l – довжина групи, l= 1.

групи	-4 …-3	-3 …-2	-2 …-1	-1 …0	0 …1	1 …2	2 …3	3 …4
m_i	6	25	72	133	120	88	46	10
p_i*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02
f_i*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02

Побудуємо графік гістограми (рис. 8.2).

З вигляду гістограми можна припустити, що її можна згладити за допомогою нормального закону (припущення підтверджується тим, що досліджувана випадкова величина є помилкою вимірювання, а отже розподілена нормально), щільність розподілу якого визначають за виразом

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}.$$

Параметри m та σ, що входять у вираз щільності f(x), оберемо так, щоб щонайкраще погодити аналітичний вираз із статистичним розподілом. Визначимо вибіркове середнє $\tilde{X}$ та вибіркове середнє квадратичне відхилення $\sigma_{выб}$ за даними групованого варіаційного ряду. Як значення x_i оберемо середину і-ї групи і цьому значенню поставимо у відповідність як імовірність його частоту p_i*, отримаємо

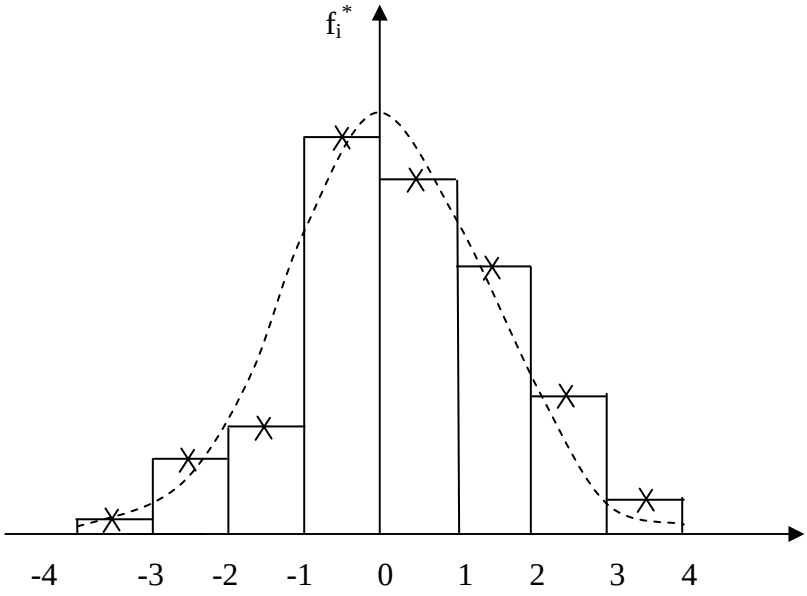

Рис. 8.2 - Графік гістограми

$$\tilde{X} = \sum_{i=1}^{8} x_i p_i^* = -3{,}5*0{,}012 - 2{,}5*0{,}05 - 1{,}5*0{,}144 - 0{,}5*0{,}266 + ;$$
$$+0{,}5*0{,}24 + 1{,}5*0{,}176 + 2{,}5*0{,}092 + 3{,}5*0{,}02 = 0{,}168$$

$$\tilde{\alpha}_2 = \sum_{i=1}^{8} x_i^2 p_i^* = (-3{,}5)^2*0{,}012 + (-2{,}5)^2*0{,}05 + (-1{,}5)^2*0{,}144 + (-0{,}5)^2*0{,}266 + ;$$
$$+(0{,}5)^2*0{,}24 + (1{,}5)^2*0{,}176 + (2{,}5)^2*0{,}092 + (3{,}5)^2*0{,}02 = 2{,}126$$

$$\sigma_{выб}^2 = \tilde{\alpha}_2 - \tilde{X}^2 = 2{,}126 - (0{,}168)^2 = 2{,}098;$$

$$\sigma_{выб} = \sqrt{\sigma_{выб}^2} = \sqrt{2{,}098} = 1{,}448.$$

Дістали розподіл $f^*(x) = \frac{1}{1{,}448\sqrt{2\pi}} \exp\left\{-\frac{(x-0{,}168)^2}{2*2{,}126}\right\}$, користуючись яким, підрахуємо значення f*(x) на межах груп:

$f_i(-4) = 0{,}0045$; $f_i(1) = 0{,}2343$;
$f_i(-3) = 0{,}0256$; $f_i(2) = 0{,}1244$;
$f_i(-2) = 0{,}0895$; $f_i(3) = 0{,}0435$;
$f_i(-1) = 0{,}1986$; $f_i(4) = 0{,}0087$;
$f_i(0) = 0{,}274$.

Відкладемо на графіку отримані точки та проведемо плавну криву (рис. 8.2).

Приклад 8.3. Для визначення точності вимірювального прибору було зроблено п'ять незалежних вимірювань, результати яких зведені у таблицю:

Номер вимірювання	1	2	3	4	5
x_i	2781	2836	2807	2763	2858

Визначити незміщену оцінку дисперсії помилок вимірювального прибору, якщо дійсне значення вимірюваної величини:

а) відомо і дорівнює 2800; б) невідомо.

Розв'язання

а) якщо значення вимірюваної величини відомо, то генеральна середня $\overline{X}$ = 2800, незміщену оцінку дисперсії у цьому випадку можна визначити за формулою

$$\sigma^2_{выб} = \frac{\sum\left(x_i - \overline{x}\right)^2}{n} = \frac{(2781-2800)^2 + (2836-2800)^2 + (2807-2800)^2 + (2763-2800)^2 + (2858-2800)^2}{5} =$$

$$= 1287{,}8\,.$$

б) якщо значення вимірюваної величини невідомо, треба визначити вибіркову середню, а незміщену оцінку дисперсії обчислити за формулою

$$S^2 = \frac{n}{n-1}\sigma^2_{выб} = \frac{\sum\left(x_i - \tilde{X}\right)^2}{n-1} = \frac{(2781-2809)^2 + (2836-2809)^2 + (2807-2809)^2 + (2763-2809)^2 + (2858-2809)^2}{5-1} =$$

$$= 1508{,}5\,.$$

Приклад 8.4. Зроблено вимірювання випадкової величини Y при різних значеннях випадкової величини X. Визначити вибірковий коефіцієнт кореляції цих величин.

x_i	-8	-10	22	2
y_i	-10	-2	4	-1

Розв'язання

Коефіцієнт кореляції визначимо за формулою

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}\,.$$

Для його обчислення необхідно знайти вибірковий кореляційний момент

$$K_{хувув} = \frac{\sum_i (x_i - \overline{X})(y_i - \overline{Y})}{n-1}\,.$$

Для оцінки середніх значень X та Y визначимо вибіркові середні

$$\tilde{X} = \frac{-8+10+22+2}{4} = 6{,}5$$

$$\tilde{Y} = \frac{-10-2+4-1}{4} = -2{,}25\,.$$

Визначимо вибіркові дисперсії X та Y

$$S_x^2 = \frac{(-8-6{,}5)^2 + (10-6{,}5)^2 + (22-6{,}5)^2 + (2-6{,}5)^2}{4-1} = 161;$$

$$S_y^2 = \frac{(-10+2{,}25)^2 + (-2+2{,}25)^2 + (4+2{,}25)^2 + (-1+2{,}25)^2}{4-1} = 33{,}6\,.$$

Визначимо вибіркові середні квадратичні відхилення:

$$\sigma_{выб_x} = \sqrt{S_x^2} = \sqrt{161} = 12{,}7 \qquad \sigma_{выб_y} = \sqrt{S_y^2} = \sqrt{33{,}6} = 5{,}8\,.$$

Тепер розрахуємо кореляційний момент

$$K_{хувув} = \frac{(-8-6{,}5)(-10+2{,}25)+(10-6{,}5)(-2+2{,}25)+(22-6{,}5)(4+2{,}25)+(2-6{,}5)(-1+2{,}25)}{4-1} = 68{,}2$$

та коефіцієнт кореляції

$$r_{xy} = \frac{K_{хувув}}{\sigma_{выб_x}\sigma_{выб_y}} = \frac{68{,}2}{12{,}7*5{,}8} = 0{,}926.$$

Приклад 8.5. Знайдемо довірчий інтервал для вибіркової середньої досліджуваної ознаки X. Нехай зроблено n незалежних дослідів та визначені спроможні і незміщені оцінки параметрів цієї ознаки $\tilde{X}$ та $\sigma^2_{виб}$.

Розв'язання

Нехай $\tilde{X}$ = 10, $\sigma^2_{виб}$ = 4, n = 40. Задамося значенням довірчої імовірності β = 0,95. Тоді можна записати

$$P\{(\tilde{X} - l) < \overline{X} < (\tilde{X} + l)\} = 0{,}95.$$

Скористуємося тим, що випадкова величина $\tilde{X}$ є функцією n незалежних випадкових величин x_i. Тоді відповідно до центральної граничної теореми щільність розподілу випадкової величини $\tilde{X}$ практично буде підпорядковуватися нормальному закону розподілу з параметрами

$$M[\tilde{X}] = \overline{X} = 10,\ D[\tilde{X}] = \frac{\sigma^2_{ген}}{n} = 4/40 = 0{,}1.$$

Для нормального закону розподілу імовірність влучення випадкової величини на інтервал значень можна виразити за допомогою інтеграла імовірностей:

$$P\{(m_x^* - l) \le m_x \le (m_x^* + l)\} = \left[\Phi\left(\frac{m_x^* + l - m_x}{\sqrt{\sigma_x}}\right) - \Phi\left(\frac{m_x^* - l - m_x}{\sqrt{\sigma_x}}\right)\right] =$$

$$= \left[\Phi\left(\frac{l}{\sqrt{\sigma_x}}\right) - \Phi\left(\frac{-l}{\sqrt{\sigma_x}}\right)\right] = 2\Phi\left(\frac{l}{\sqrt{\sigma_x}}\right) = \beta\,.$$

Підставимо значення:

$$\Phi\left(\frac{l}{\sqrt{0{,}32}}\right) = 0{,}475; \qquad \frac{l}{0{,}453} = 1{,}4\,,$$

звідки l = 1,4 * 0,453 = 0,634.

Отже, з імовірністю 0,95 інтервал (9,366; 10,634) накриє генеральну середню досліджуваної ознаки X.

Приклад 8.6 Результати вимірювання значень випадкових величин X та Y наведені у таблиці:

№ досліду	**x_i**	**y_i**
1	**2**	**3**
1	4	0,041
2	8	0,05
3	10	0,081

Продовження

1	2	3
4	14	0,104
5	16	0,12
6	20	0,139
7	19	0,154
8	23	0,18
9	26	0,208
10	30	0,241
11	31	0,25
12	36	0,269
13	37	0,301

Визначити точкові й інтервальні оцінки числових характеристик системи випадкових величин X та Y, а також імовірність того, що вибіркова середня випадкової величини X відрізняється від її генеральної середньої не більш чим на 1.

Розв'язання

Проміжні розрахунки будемо зводити у таблицю. Спочатку визначимо точкові значення вибіркових середніх:

$$\tilde{Y} = \frac{\sum y_i}{n} = \frac{2{,}138}{13} = 0{,}164;$$

$$\tilde{X} = \frac{\sum x_i}{n} = \frac{274}{13} = 21{,}1.$$

№ досліду	x_i	y_i	$\overset{o}{x}_i$	$\overset{o}{x}^2$	$\overset{o}{y}_i$	$\overset{o}{y}^2$	$\overset{o}{x}_i \overset{o}{y}_i$
1	4	0,041	-17,1	292,41	-0,123	0,0151	2,103
2	8	0,05	-13,1	171,61	-0,114	0,0130	1,493
3	10	0,081	-11,1	123,21	-0,083	0,0069	0,921
4	14	0,104	-7,1	50,41	-0,06	0,0036	0,426
5	16	0,12	-5,1	26,01	-0,044	0,0019	0,224
6	20	0,139	-1,1	1,21	-0,025	0,0006	0,028
7	19	0,154	-2,1	4,41	-0,01	0,0001	0,021
8	23	0,18	1,9	3,61	0,016	0,0003	0,030
9	26	0,208	4,9	24,01	0,044	0,0019	0,216
10	30	0,241	8,9	79,21	0,077	0,0059	0,685
11	31	0,25	9,9	98,01	0,086	0,0074	0,851
12	36	0,269	14,9	222,01	0,105	0,0110	1,565
13	37	0,301	15,9	252,81	0,137	0,0188	2,178
Сума	$\Sigma x_i=274$	$\Sigma y_i=2{,}138$		$\sum \overset{o}{x}_i^2$		$\sum \overset{o}{y}_i^2$	$\sum \overset{o}{x}_i \overset{o}{y}_i$

Знайдемо точкові значення незміщених вибіркових дисперсій:

$$S_y^2 = \frac{\sum_1^{13}(y_i - \tilde{Y})^2}{13-1} = \frac{\sum_1^{13}\overset{o}{y}^2}{12} = \frac{0{,}0866}{12} = 0{,}0072;$$

$$S_x^2 = \frac{\sum_1^{13}(x_i - \tilde{X})^2}{13-1} = \frac{\sum_1^{13}\overset{o}{x}^2}{12} = \frac{1348{,}93}{12} = 112{,}41.$$

Визначимо вибіркові середні квадратичні відхилення

$$\sigma_{виб_y} = \sqrt{0{,}0072} = 0{,}085; \quad \sigma_{виб_x} = \sqrt{112{,}41} = 10{,}6.$$

Визначимо вибірковий кореляційний момент за формулою

$$K_{хувув} = \frac{\sum_1^{13}\overset{o}{x}\,\overset{o}{y}}{13-1} = \frac{10{,}742}{12} = 0{,}895$$

та коефіцієнт кореляції за формулою

$$r_{хувув} = \frac{K_{хувув}}{\sigma_{виб_y} * \sigma_{виб_x}} = \frac{0{,}895}{0{,}085 * 10{,}6} = 0{,}994.$$

Для визначення імовірності того, що помилка від заміни генеральної середньої випадкової величини X її вибірковою оцінкою не перевершить 1, побудуємо довірчий інтервал з межами 21,1±1 та визначимо імовірність того, що цей інтервал накриє генеральну середню випадкової величини X (довірчу імовірність β).

Скористуємося тим, що величина $\tilde{X}$ є сумою n=13 незалежних однаково розподілених випадкових величин x_i і, відповідно до центральної граничної теореми, при досить великому n її закон розподілу близький до нормального (а в нашому випадку ми провадили вимірювання, що завжди дає помилку, розподілену за нормальним законом). Будемо виходити з того, що величина $\tilde{X}$ розподілена за нормальним законом з параметрами:

$$M[\tilde{X}] = \overline{X} = 21{,}1 \quad \text{та} \quad \sigma[\tilde{X}] = \sqrt{\frac{\sigma_{ген}^2}{n}} = \frac{10{,}6}{\sqrt{13}} = 2{,}94.$$

Виразимо шукану імовірність β за допомогою функції Лапласа:

$$P\left\{\left|\tilde{X} - \overline{X}\right| < l\right\} = \beta = 2\Phi\left(\frac{l}{\sigma[\tilde{X}]}\right),$$

звідки дістанемо

$$2\Phi\left(\frac{1}{2{,}94}\right) = \beta;\ \beta = 2\Phi(0{,}34) = 0{,}266.$$

Отримане значення імовірності дуже мале, тобто подія, що полягає в тому, що помилка від заміни генеральної середньої X її вибірковою оцінкою не перевершить 1, практично неможлива. Задамося довірчою ймовірністю β =0,95

та визначимо межі відповідного їй довірчого інтервалу. Для цього запишемо вираз імовірності β

$$2\Phi\left(\frac{l}{2{,}94}\right) = \beta = 0{,}95;\ \Phi\left(\frac{l}{2{,}94}\right) = 0{,}95/2 = 0{,}475,$$

звідки, користуючись таблицею значень інтеграла імовірностей, дістанемо

$$\frac{l}{2{,}94} = 1{,}96,$$

отже, l=1,96*2,94 = 5,76. З імовірністю 0,95 інтервал (15,34; 26,86) накриє генеральну середню випадкової величини X. Помилка становить 27,3%, тобто занадто велика. Для зменшення помилки необхідно збільшити кількість вимірювань.

Побудуємо довірчий інтервал для вибіркової дисперсії випадкової величини X. Закон розподілу вибіркової дисперсії також наближається до нормального. Один з параметрів розподілу - математичне сподівання вибіркової дисперсії

$$M[S^2_x] = \sigma^2_{ген.}$$

Дисперсію вибіркової дисперсії S^2_x обчислимо за формулою:

$$D[S_x^2] = \frac{1}{n}\mu_4 - \frac{n-3}{n(n-1)}\sigma^2_{ген},$$

де μ_4 – центральний момент випадкової величини X четвертого порядку, його оцінку обчислимо за формулою:

$$\mu_4^* = \frac{1}{n}\sum_i (x_i - \tilde{X})^4.$$

Але справа в тому, що при обмеженій кількості дослідів моменти високого порядку визначаються з великими помилками, тому якщо величина X розподілена за нормальним законом, то μ_4 можна обчислити через дисперсію:

$$\mu_4 = 3*\sigma^2_{ген}.$$

Отримаємо:

$$D[S_x^2] = \frac{3}{n}\sigma^2_{ген} - \frac{n-3}{n(n-1)}\sigma^2_{ген} = \frac{3}{13}(114{,}041)^2 - \frac{13-3}{13(13-1)}(114{,}041)^2 = 2106,$$

звідси

$$\sigma[S_x^2] = \sqrt{2106} = 45{,}9.$$

Визначимо довірчий інтервал для вибіркової дисперсії, задавшись імовірністю β = 0,9.

Користуючись тим, що S^2_x розподілена нормально, виразимо імовірність β за допомогою функції Лапласа:

$$P\left\{\left|S_x^2 - D_x\right| < l\right\} = \beta = 2\Phi\left(\frac{l}{\sigma[S_x^2]}\right),\ \text{звідки дістанемо}\ \ \Phi\left(\frac{l}{45{,}9}\right) = 0{,}45,$$

$$\frac{l}{45{,}9} = 1{,}64, \qquad l=1{,}64*45{,}9 = 75{,}26.$$

Отже, з імовірністю 0,9 дисперсія X лежить в інтервалі 112,41±75,26. Помилка становить 67% - дуже велика.

Приклад 8.7. У попередньому прикладі 6.4 результати обчислення показали, що для зменшення довірчого інтервалу (15,34; 26,86), який з довірчою імовірністю 0,95 накриє генеральну середню випадкової величини X, необхідно

збільшити кількість вимірювань. Визначимо, якою повинна бути кількість вимірювань n, щоб помилка від заміни генеральної середньої випадкової величини X її вибірковою оцінкою $\tilde{X} = 21{,}1$ не перевершувала 1 з імовірністю β=0,95.

Розв'язання

Виразимо довірчу імовірність β за допомогою функції Лапласа:

$$P\left\{\left|\tilde{X} - \overline{X}\right| < l\right\} = \beta = 2\Phi\left(\frac{l}{\sigma[\tilde{X}]}\right),$$

звідки дістанемо, підставивши l=1 та β =0,95,

$$\Phi\left(\frac{1}{\sigma[\tilde{X}]}\right) = 0{,}475,$$

тоді $\frac{1}{\sigma[\tilde{X}]} = 1{,}96$, звідки $\sigma[\tilde{X}] = \frac{1}{1{,}96} = 0{,}510$. Відомо, що середнє квадратичне відхилення вибіркової середньої $\sigma[\tilde{X}] = \sqrt{\frac{\sigma^2_{ген}}{n}}$. Визначимо необхідне число вимірів n у такий спосіб:

$$n = \left(\frac{\sigma^2_{ген}}{\sigma[\tilde{X}]}\right)^2 = \left(\frac{10{,}6^2}{0{,}510}\right)^2 = 432.$$

Отже, для досягнення необхідної точності вибіркової середньої замість 13 дослідів необхідно зробити 432 вимірювання випадкової величини X.

Задачі для самостійного розв'язання:

8.1. На біржі протягом деякого часу проводяться статистичні дослідження коливань ціни на партії товарів. Результати спостережень: 3100, 4000, 3800, 4100, 3400, 4200, 3700, 3900, 3200, 4100, 3800, 4200, 3500, 4000, 3900. Побудувати варіаційний ряд і гістограму випадкової величини X - ціна на партію товарів.

8.2. Для умов попередньої задачі обчислити числові характеристики варіаційного ряду.

8.3. Для умов задач 8.1 і 8.2 визначити довірчі інтервали для вибіркової середньої з надійністю 0,95 і вибіркової дисперсії з надійністю 0,99. Прийняти, що досліджувана ознака X розподілена нормально.

Тема 9. ЕЛЕМЕНТИ ТЕОРІЇ КОРЕЛЯЦІЇ (14 годин)

Функціональна, статистична та кореляційна залежності. Рівняння регресії. Поле кореляції. Коефіцієнт регресії. Метод найменших квадратів. Вибірковий коефіцієнт кореляції. Вибіркове кореляційне відношення. Міжгрупова та внутрігрупова дисперсії.

Література: [1] с. 249-260; 271-279; [3] с. 108-131; [4] с. 445-450; [5] с. 365-381;.

Запитання для самоперевірки:

1. Які задачі вирішують методом кореляційного аналізу?
2. В яких випадках залежність y = f(x) є функціональною, статистичною або кореляційною?
3. Дайте визначення термінів «регресія», «лінія регресії», «рівняння регресії».
4. Поясніть значення термінів «пояснююча змінна», «результативна ознака».

5. З яких міркувань визначають тип кореляційної залежності y= f(x)? Які типи залежностей Ви знаєте?

6. Чим характерна лінійна залежність y = f(x)? Чому її використовують найчастіше?

7. Як називаються параметри лінійної залежності y = f(x)?

8. Які методи можна використовувати для визначення параметрів рівняння регресії y = f(x)?

9. Якій вимозі задовольняють параметри, визначені за методом найменших квадратів?

10. Назвіть характеристики, що дозволяють оцінити наявність зв'язку між ознакою-фактором і результативною ознакою.

11. Які значення може приймати коефіцієнт кореляції, які висновки можна зробити на підставі цих значень?

12. Які значення може приймати кореляційне відношення, і які висновки можна зробити на підставі цих значень?

13. Що таке кореляційна таблиця?

14. Які параметри визначають за допомогою кореляційної таблиці?

Приклад 9.1. Нехай у результаті досліду отримані наступні експериментальні дані:

x_i	1	2	3
y_i	1	3	4

Потрібно визначити параметри лінійної та квадратичної залежностей для X та Y.

Розв'язання

Нанесемо на координатну площину точки з координатами (x_i,y_i) (рис. 9.1).

Із графіка видно, що точки не лежать на одній прямій, і що із зростанням X У має тенденцію до зростання.

а) нехай шукана залежність - лінійна:

$$y = a_0 + a_1 x,$$

запишемо її як функцію параметрів: $\sum_{i=1}^{n}(y_i - a_0 - a_1 x_i)^2 \to min$ та візьмемо часткові похідні за параметрами a_1 та a_0 і дорівняємо їх до нуля:

$$\begin{cases} 2*\Sigma(y_i - a_0 - a_1 x_i)*(-1) = 0; \\ 2*\Sigma(y_i - a_0 - a_1 x_i)*(-x_i) = 0; \end{cases}$$

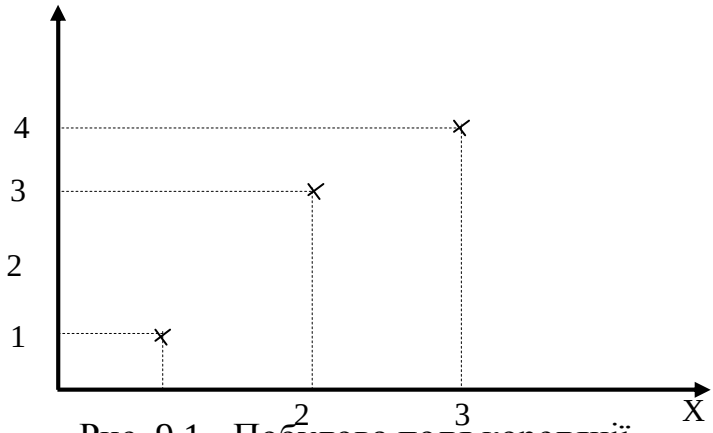

Рис. 9.1 - Побудова поля кореляції

виконаємо перетворення:

$$\begin{cases} \Sigma a_0 + \Sigma a_1 x_i = \Sigma y_i \\ \Sigma a_0\, x_i + \Sigma a_1 x_i^2 = \Sigma y_i\, x_i \end{cases}$$

$$\begin{cases} n a_0 + a_1 \Sigma x_i = \Sigma y_i \\ a_0 \Sigma\, x_i + a_1 \Sigma x_i^2 = \Sigma y_i\, x_i \end{cases}$$

Підставимо значення: $n=3$; $\Sigma x_i = 6$; $\Sigma x_i^2 = 14$; $\Sigma y_i = 8$; $\Sigma y_i\, x_i = 19$ та визначимо параметри.

$$\begin{cases} 3a_0 + 6\, a_1 = 8 \\ 6a_0 + 14a_1 = 19. \end{cases}$$

$$\begin{cases} a_0 = -1/3 \\ a_1 = 1{,}5. \end{cases}$$

Шукана залежність має вигляд:

$$y = -1/3 + 1{,}5\, x.$$

Отримана лінійна залежність є найімовірнішою з лінійних залежностей.

Протабулюємо її та побудуємо графік (рис. 9.2).

x_i	1	2	3
y_i	1	3	4
$y_i^т$	1,17	2,67	4,17

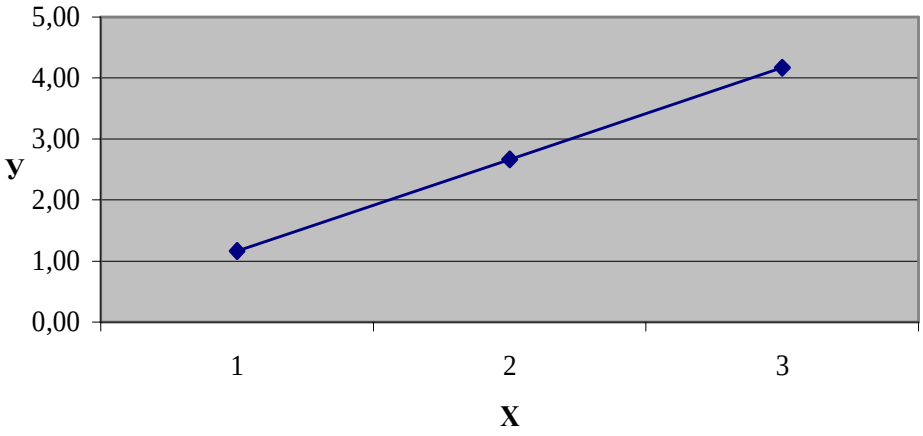

Рис. 9.2 - Лінійна залежність y = -1/3 + 1,5 x.

б) нехай шукана залежність – квадратична: $Y = a_0 + a_1x + a_2x^2$, запишемо-мо її: $\sum_{i=1}^{n}(y_i - a_0 - a_1x_i - a_2x_i^2)^2 \to min$, візьмемо часткові похідні та дорівняємо їх до нуля:

$$\begin{cases} 2*\Sigma(y_i - a_0 - a_1x_i - a_2x_i^2)*(-1) = 0 \\ 2*\Sigma(y_i - a_0 - a_1x_i - a_2x_i^2)*(-x_i) = 0 \\ 2*\Sigma(y_i - a_0 - a_1x_i - a_2x_i^2)*(-x_i^2) = 0, \end{cases}$$

виконаємо перетворення:

$$\begin{cases} \Sigma y_i - \Sigma a_0 - \Sigma a_1x_i - \Sigma a_2x_i^2 = 0 \\ \Sigma y_i\, x_i - \Sigma\, a_0\, x_i - \Sigma a_1x_i^2 - \Sigma a_2x_i^3 = 0 \\ \Sigma y_i\, x_i^2 - \Sigma a_0\, x_i^2 - \Sigma a_1x_i^3 - \Sigma a_2x_i^4 = 0 \end{cases}$$

$$\begin{cases} na_0 + a_1\Sigma x_i + a_2\Sigma x_i^2 = \Sigma y_i \\ a_0\, \Sigma x_i + a_1\Sigma x_i^2 + a_2\Sigma x_i^3 = \Sigma y_i\, x_i \\ a_0\, \Sigma x_i^2 + a_1\Sigma x_i^3 + a_2\Sigma x_i^4 = \Sigma y_i\, x_i^2 \end{cases}$$

Підставимо значення: $n=3$; $\Sigma x_i = 6$; $\Sigma x_i^2 = 14$; $\Sigma y_i = 8$; $\Sigma y_i\, x_i = 19$; $\Sigma x_i^3 = 36$; $\Sigma x_i^4 = 98$; $\Sigma y_i\, x_i^2 = 49$ та визначимо параметри

$$\begin{cases} 3a_0 + 6\, a_1 + 14a_2 = 8 \\ 6a_0 + 14a_1 + 36a_2 = 19 \\ 14a_0 + 36a_1 + 98a_2 = 49. \end{cases}$$

Звідки : $a_2=-0{,}091$; $a_1=1{,}273$; $a_0=0{,}0455$.

Отже, найімовірніша квадратична залежність матиме вигляд:

$$y = 0{,}0455 + 1{,}273x - 0{,}091x^2.$$

Протабулюємо її, збільшивши для наочності кількість точок:

x_i	0	1	2	3	4	5
$y_i^т$	0,05	1,23	2,23	3,05	3,68	4,14

Отримаємо графік, показаний на рис. 9.3.

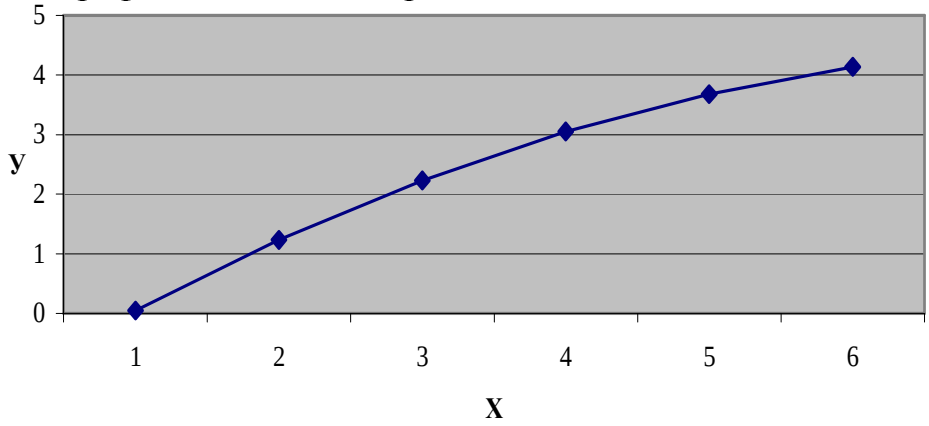

Рис. 9.3 - Квадратична залежність $y = 0{,}0455 + 1{,}273x - 0{,}091x^2$

Приклад 9.2. Визначити параметри лінійної регресії для системи випадкових величин X і Y, результати вимірювання яких приведені у задачі 9.1.

№ досліду	x_i	y_i
1	4	0,041
2	8	0,05
3	10	0,081
4	14	0,104
5	16	0,12
6	20	0,139
7	19	0,154
8	23	0,18
9	26	0,208
10	30	0,241
11	31	0,25
12	36	0,269
13	37	0,301

Розв'язання

Побудуємо поле кореляції (рис. 9.4). Очевидно, що статистичні дані добре згладжуються лінійною залежністю.

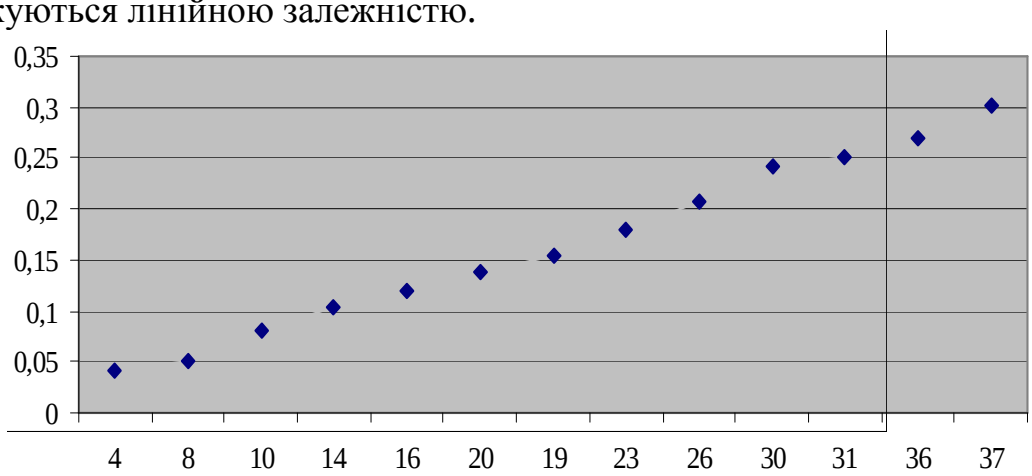

Рис. 9.4 - Побудова поля кореляції

Визначимо параметри лінійної залежності між X та Y, проміжні обчислення зробимо у таблиці

№ досліду	$\mathbf{x_i}$	$\mathbf{y_i}$	$\mathbf{x_i^2}$	$\mathbf{x_iy_i}$
1	**2**	**3**	**4**	**5**
1	4	0,041	16	0,164
2	8	0,05	64	0,4
3	10	0,081	100	0,81
4	14	0,104	196	1,456
5	16	0,12	256	1,92

Продовження таблиці

1	2	3	4	5
6	20	0,139	400	2,78
7	19	0,154	361	2,926
8	23	0,18	529	4,14
9	26	0,208	676	5,408
10	30	0,241	900	7,23
11	31	0,25	961	7,75
12	36	0,269	1296	9,684
13	37	0,301	1369	11,137
Сума	Σx_i=274	Σy_i=2,138	Σx_i^2= 7124	$\Sigma x_i * y_i$= 55,805

Отримаємо

$$\rho_{yx} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{13*55{,}08 - 274*2{,}138}{13*7124 - 274^2} = 0{,}00796\,;$$

$$b = \frac{\sum x_i^2 * \sum y_i - \sum x_i * \sum x_i y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{7124*2{,}138 - 274*55{,}8}{13*7124 - 274^2} = -0{,}0034\,.$$

Отже, шукана залежність матиме вигляд:

$$y = 0{,}00796x - 0{,}0034.$$

Приклад 9.3. Визначимо значення коефіцієнта кореляції системи випадкових величин X та Y з задачі 6.4 та задачі 9.2.

Розв'язання

У результаті розв’язання попередніх задач для системи випадкових величин X та Y отримано лінійну кореляційну залежність. Скористаємося параметрами цієї залежності та визначимо коефіцієнт кореляції, розрахувавши попередньо середні квадратичні відхилення σ_x та σ_y:

$$\sigma_x = 10{,}6;\ \sigma_y = 0{,}085;$$

$$r_{вху} = \rho * \frac{\sigma_x}{\sigma_y} = 0{,}00796\frac{10{,}6}{0{,}085} = 0{,}994\,.$$

Приклад 9.4. Визначити значення коефіцієнта кореляції системи випадкових величин I та U. Кореляційна залежність, що отримана за результатами 20 дослідів має вигляд: u = 1,525i + 3,948, середні квадратичні відхилення σ_i та σ_u.відповідно дорівнюють σ_i = 12,72; σ_u = 23,64.

Розв'язання

Розрахуємо і визначимо коефіцієнт кореляції за формулою:

$$r_{віu} = \rho * \frac{\sigma_i}{\sigma_u} = 1{,}525\frac{12{,}72}{23{,}64} = 0{,}821\,.$$

Отже, між I та U існує тісний лінійний зв'язок.

Задачі для самостійного розв'язання:

9.1. Результати вимірювань досліджуваної ознаки Y зведені в таблицю

x_i	41	50	81	104	120	139	154	180	208	241	250	269	301
y_i	4	8	10	14	16	20	14	23	26	30	31	36	37

Використовуючи поле кореляції, вибрати клас залежності y=((x), побудувати рівняння регресії Y на X, оцінити тісноту зв'язку між фактором і результативною ознакою.

9.2. Результати вимірювань фактору X і результативної ознаки Y наведені в таблиці

x_i	1	2	3	4	5	6	7	8
y_i	6	17	34	57	86	121	162	209

Користуючись методом найменших квадратів визначити параметри апроксимуючої залежності $y=ax^2+bx+c$.

9.3. У результаті статистичних спостережень зареєстрована залежність u=f(t).

u_i	75	55	40	30	20	15	10	10	5	5
t_i	1	2	3	4	5	6	7	8	9	10

Визначити параметри експонентної апроксимуючої залежності.

Тема 10. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ (14 годин)

Поняття статистичної гіпотези. Нульова та альтернативна гіпотези. Критична область. Область прийняття гіпотези. Критична точка zk. Статистичний критерій. Помилки 1-го та 2-го роду. Потужність критерію. Порівняння вибіркової середньої та генеральної середньої нормальної сукупності. t-критерій (розподіл Стьюдента). Порівняння двох дисперсій нормальних генеральних сукупностей. F-критерій (розподіл Фішера). Порівняння вибіркової та генеральної дисперсій нормальної сукупності. Критерій $\chi 2$. Перевірка гіпотези про нормальний розподіл генеральної сукупності. Елементи дисперсійного аналізу. Загальна, факторна та залишкова сума квадратів відхилень. Загальна, факторна та залишкова дисперсії. Елементи регресійного аналізу.

Література: [1] С. 282-290; 332-338; 342-351; [3] С. 132-162; [4] С. 445-450; 467-470; [5] С. 281-287.

Запитання для самоперевірки:

1. Поясніть, що таке статистична гіпотеза?
2. Поясніть, що таке нульова і альтернативна гіпотези?
3. Які критерії застосовують для перевірки статистичних гіпотез?
4. Поясніть значення термінів «критична область», «критична точка», «область прийняття гіпотези».

5. В якому випадку слід вибирати двосторонню критичну область?

6. Який результат перевірки гіпотези відносять до помилки 1-го роду та який до помилки 2-го роду?

7. Наведіть приклади задач на перевірку гіпотез.

8. Які задачі вирішують за допомогою однофакторного дисперсійного аналізу?

9. Чому цей вид аналізу випадкових даних отримав назву дисперсійного?

10. Що являють собою величини $S_{общ}$, $S_{факт}$ і $S_{ост}$? Яке співвідношення між ними?

11. Які задачі вирішують методом регресійного аналізу?

12. Для чого застосовують метод найменших квадратів?

13. Що таке рівняння регресії? Як перевіряють його адекватність статистичним даним?

14. Як визначають значущість коефіцієнтів рівняння регресії?

Задачі для самостійного розв'язання:

10.1. Число появ герба при 20 підкиданнях двох монет розподілилося в такий спосіб:

Число гербів	0	1	2
Число появ	4	8	8

Чи узгоджуються ці дані з припущенням про симетричність монет і незалежність результатів підкидання. Як рівень значущості прийняти α=0,05.

ЗМ 3. ВИПАДКОВІ ПРОЦЕСИ

Тема 11. ЕЛЕМЕНТИ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ (16 годин)

Поняття випадкового процесу. Імовірнісні характеристики випадкового процесу. Кореляційна функція та її властивості. Лінійне перетворення суми випадкових функцій Xi(t). Стаціонарний випадковий процес. Математичне сподівання та дисперсія стаціонарного випадкового процесу. Кореляційна функція стаціонарного випадкового процесу. Властивість ергодичності. Числові характеристики, визначені за множиною реалізацій та за часом. Елементи теорії масового обслуговування. Основні поняття. Класифікація СМО. Вхідний потік вимог. Дисципліна черги. Механізм обслуговування. Поняття марківського випадкового процесу. Випадковий процес з дискретними станами та безперервним часом. Граф станів системи. Рівняння для імовірностей станів. Одноканальна СМО з відмовами. Фінальні імовірності станів. Характеристики системи масового обслуговування. Багатоканальна СМО з відмовами. Одноканальна СМО з очікуванням та обмеженою чергою. Одноканальна СМО з очікуванням та необмеженою чергою. Багатоканальна СМО з очікуванням та необмеженою чергою. Комп’ютерна обробка статистичних даних.

Література: [3] С. 163-203; [5] С. 386-409; 412-421.

Запитання для самоперевірки:

1. Дайте визначення випадкового процесу.
2. Що таке реалізація випадкової функції.

3. Які властивості імовірнісних характеристик стаціонарного випадкового процесу?

4. Дайте визначення кореляційної функції.

5. Поясніть властивість ергодичності стаціонарного випадкового процесу.

6. Які задачі вирішує теорія масового обслуговування?

7. Перелічіть основні компоненти системи масового обслуговування, що обумовлюють її математичний опис.

8. Поясніть класифікацію систем масового обслуговування.

9. В якому випадку випадковий процес, що протікає в системі, називається марковським?

10. Дайте визначення стаціонарного потоку подій і регулярного потоку подій.

11. Які змінні є невідомими в рівняннях Колмогорова? Що таке фінальної імовірності?

12. Якими параметрами визначається потік заявок? Потік обслуговувань?

13. В чому полягає суть методу статистичних випробувань?

14. В чому полягає суть моделювання у нейронних мережах?

Задачі для самостійного розв'язання:

9.1. Випадкова функція X(t) в кожному перетині являє собою безперервну випадкову величину з щільністю розподілу f(x,t). Напишіть вираз для математичного сподівання $m_x(t)$ і дисперсії $D_x(t)$ випадкової функції X(t).

9.2. Випадкова функція X(t) задана у вигляді X(t) = Vt+1, де V – випадкова величина, розподілена за нормальним законом з параметрами $m_v=1$ і $\sigma_v=0,2$. Визначити числові характеристики випадкової функції X(t).

9.3. Розглядається робота автоматичної телефонної станції, розрахованої на одночасне обслуговування 20 абонентів. Виклик на АТС надходить в середньому через 6 секунд. Кожна розмова триває в середньому 2 хвилини. Якщо абонент застає АТС зайнятою, він отримує відмову. Якщо абонент застає вільним хоча б один з 20 каналів, він з'єднується з потрібним йому номером. Визначити імовірність того, що абонент буде обслужений, а також інші характеристики роботи СМО: середнє число зайнятих каналів, імовірність зайнятості каналу, середній час простою каналу.

СПИСОК ДЖЕРЕЛ

1. Гмурман В. Э. Теория вероятностей и математическая статистика. – М.: Высш. школа, 1977. – 498 с.
2. Гмурман В.Э. Руководство к решению задач по теории вероятностей и математической статистике. - М.: Высш. школа, 1975. – 330 с.
3. Ачкасов А.Є., Плакіда В.Т., Воронков О.О., Воронкова Т.Б. Теорія імовірностей і математична статистика: Навчальний посібник. – Харків, ХНАМГ, 2008. – 249 с.
4. Вентцель Е. С. Теория вероятностей. – М.: Высшая школа, 1999.
5. 14Теория статистики с основами теории вероятностей: Учебное пособие для вузов/ Под ред. И.И. Елисеевой. – М.: ЮНИТИ-ДАНА, 2001. – 446 с.
6. Карасев А.И., Аксютина З.И., Савельева Т.И. Курс высшей математики для экономических вузов. ч.2. – М.:Высш. школа, 1982. – 320 с.
7. Справочник по математике для экономистов (Под редакцией В.И.Ермакова.) – М.:Высш.школа, 1987. – 306 с.
8. http://teorver.ru
9. http://www.artspb.com
10. http://www.matburo.ru
11. http://stud-project.ru
12. http://www.statsoft.ru
13. http://www.alife.narod.ru
14. http://neuro.net.ua

ДОДАТКИ

Додаток 1

Таблиця значень функції $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0008	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Додаток 2.

Таблиця значень функції $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

X	Φ(x)	X	Φ(x)	X	Φ(x)	X	Φ(x)
0,00	0,0000	0,36	0,1406	0,72	0,2642	1,08	0,3599
0,01	0,0040	0,37	0,1443	0,73	0,2673	1,09	0,3621
0,02	0,0080	0,38	0,1480	0,74	0,2703	1,10	0,3643
0,03	0,0120	0,39	0,1517	0,75	0,2734	1,11	0,3665
0,04	0,0160	0,40	0,1554	0,76	0,2764	1,12	0,3686
0,05	0,0199	0,41	0,1591	0,77	0,2794	1,13	0,3708
0,06	0,0239	0,42	0,1628	0,78	0,2823	1,14	0,3729
0,07	0,0279	0,43	0,1664	0,79	0,2852	1,15	0,3749
0,08	0,0319	0,44	0,1700	0,80	0,2818	1,16	0,3770
0,09	0,0359	0,45	0,1736	0,81	0,2910	1,17	0,3790
0,10	0,0398	0.46	0,1772	0,82	0,2939	1,18	0,3810
0,11	0,0438	0,47	0,1808	0,83	0,2967	1,19	0,3830
0,12	0,0478	0,48	0,1844	0,84	0,2995	1,20	0,3849
0,13	0,0517	0,49	0,1879	0,85	0,3023	1.21	0,3869
0,14	0,0557	0,50	0,1915	0,86	0,3051	1,22	0,3883
0,15	0,0596	0,51	0,1950	0,87	0,3078	1,23	0,3907
0,16	0,0636	0,52	0,1985	0,88	0,3106	1,24	0,3925
0,17	0,0675	0,53	0,2019	0,89	0,3133	1,25	0,3944
0,18	0,0714	0,54	0,2054	0,90	0,3159	1,26	0,3926
0,19	0,0753	0,55	0,2088	0,91	0,3186	1,27	0,3980
0,20	0,0793	0,56	0,2123	0,92	0,3212	1,28	0,3997
0,21	0,0832	0,57	0,2157	0,93	0,3238	1,29	0,4015
0,22	0,0871	0,58	0,2190	0,94	0,3264	1,30	0,4032
0,23	0,0910	0,59	0,2224	0,95	0,3289	1,31	0,4049
0,24	0,0948	0,60	0,2257	0,96	0,3315	1,32	0,4066
0,25	0,0987	0,61	0,2291	0,97	0,3340	1,33	0,4082
0,26	0,1026	0,62	0,2324	0,98	0,3365	1,34	0,4099
0.27	0,1064	0,63	0,2357	0,99	0,3389	1,35	0,4115
0,28	0,1103	0,64	0,2389	1,00	0,3413	1,36	0,4131
0,29	0,1141	0,65	0,2422	1,01	0,3438	1,37	0,4147
0,30	0,1179	0,66	0,2454	1,02	0,3461	1,38	0,4162
0,31	0,1217	0,67	0,2486	1,03	0,3485	1,39	0,4177
0,32	0,1255	0,68	0,2517	1,04	0,3508	1,40	0,4192
0,33	0,1293	0,69	0,2549	1,05	0,3531	1,41	0,4207
0,34	0,1331	0,70	0,2580	1,06	0,3554	1,42	0,4222
0,35	0,1368	0,71	0,2611	1,07	0,3577	1,43	0,4236

Продовження додатка 2

x	*Ф(x)*	*X*	*Ф(x)*	*X*	*Ф(x)*	*X*	*Ф(x)*
1,44	0,4251	1,73	0,4582	2,04	0,4793	2,62	0,4956
1,45	0,4265	1,74	0,4591	2,06	0,4803	2,64	0,4959
1,46	0,4279	1,75	0,4599	2,08	0,4812	2,66	0,4961
1,47	0,4292	1,76	0,4608	2,10	0,4821	2,68	0,4963
1,48	0,4306	1,77	0,4616	2,12	0,4830	2,70	0,4965
1,49	0,4319	1,78	0,4625	2,14	0,4838	2,72	0,4967
1,50	0,4332	1,79	0,4633	2,16	0,4846	2,74	0,4969
1,51	0,4345	1,80	0,4641	2,18	0,4854	2,76	0,4971
1,52	0,4357	1,81	0,4649	2,20	0,4861	2,78	0,4973
1,53	0,4370	1,82	0,4656	2,22	0,4868	2,80	0,4974
1,54	0,4382	1,83	0,4664	2,24	0,4875	2,82	0,4976
1,55	0,4394	1,84	0,4671	2,26	0,4881	2,84	0,4977
1,56	0,4406	1,85	0,4678	2,28	0,4887	2,86	0,4979
1,57	0,4418	1,86	0,4686	2,30	0,4893	2,88	0,4980
1,58	0,4429	1,87	0,4693	2,32	0,4898	2,90	0,4981
1,59	0,4441	1,88	0,4699	2,34	0,4904	2,92	0,4982
1,60	0,4452	1,89	0,4706	2,36	0,4909	2,94	0,4984
1,61	0,4463	1,90	0,4713	2,38	0,4913	2,96	0,4985
1,62	0,4474	1,91	0,4719	2,40	0,4918	2,98	0,4986
1,63	0,4484	1,92	0,4726	2,42	0,4922	3,00	0,49865
1,64	0,4495	1,93	0,4732	2,44	0,4927	3,20	0,49931
1,65	0,4505	1,94	0,4738	2,46	0,4931	3,40	0,49966
1,66	0,4515	1,95	0,4744	2,48	0,4934	3,60	0,499841
1,67	0,4525	1,96	0,4750	2,50	0,4938	3,80	0,499928
1,68	0,4535	1,97	0,4756	2,52	0,4941	4,00	0,499968
1,69	0,4545	1,98	0,4761	2,54	0,4945	4,50	0,499997
1,70	0,4554	1,99	0,4767	2,56	0,4948	5,00	0,499997
1,71	0,4564	2,00	0,4772	2,58	0,4951		
1,72	0,4573	2,02	0,4783	2,60	0,4953		

Навчальне видання

МЕТОДИЧНІ ВКАЗІВКИ

до самостійної роботи
з курсу

ТЕОРІЯ ІМОВІРНОСТЕЙ

(для студентів 2 курсу заочної форми навчання ФПО та ЗН напряму підготовки 6.060101 «Будівництво» спеціальності «Міське будівництво і господарство»)

Укладачі : доц. **ОХРІМЕНКО** Вячеслав Миколайович,
ст. викл. **ВОРОНКОВА** Тетяна Борисівна,
ст. викл. **ВОРОНКОВ** Олексій Олександрович.

Відповідальний за випуск*: А. С. Гаєвський*

За авторською редакцією

Комп'ютерне верстання *І. В. Волосожарова*

План 2012, поз. 611М

Підп. до друку 21.06.2012 р. Формат 60х84/16
Друк на ризографі Ум. друк. арк. 2,4
Тираж 50 пр. Зам. №

Видавець і виготовлювач:
Харківський національний університет
міського господарства імені О. М. Бекетова,
вул. Революції, 12, Харків, 61002
Електронна адреса: rectorat@kname.edu.ua
Свідоцтво суб'єкта видавничої справи:
ДК №4705 від 28.03.2014 р.