# ВИВЧЕННЯ ОСНОВНИХ ПОНЯТЬ ТЕОРІЇ ІНФОРМАЦІЇ У КУРСІ ТЕОРІЇ ЙМОВІРНОСТЕЙ

***Чепорнюк І.Д.***
*старший викладач*
*Київська академія водного транспорту*
*імені гетьмана П.Конашевича-Сагайдачного*

Обгрунтовано доцільність вивчення елементів теорії інформації в курсі теорії ймовірностей та математичної статистики студентами спеціальності „Програмування електронно-обчислювальної техніки та автоматизованих систем", досліджено місце, зміст та методичні особливості вивчення основних понять теорії інформації.

Обоснована целесообразность изучения элементов теории информации в курсе теории вероятностей и математической статистики студентами специальности «Программирование электронно-вычислительной техники и автоматизированных систем», исследованы место, содержание иметодическите особенности изучения основных понятий теории информации.

The necessity of study of information theory in the course of probability theory and mathematical statistics students specializing in programming electronic computers and computer systems, space research, content and methodological features of the study of basic concepts of information theory.

## Вступ

Теорія інформації – це наука, що вивчає кількісні закономірності, пов'язані з отриманням, передаванням, обробкою та зберіганням інформації. Ця теорія виникла з практичних задач теорії зв'язку і на даний час є необхідним математичним апаратом для вивчення всеможливих процесів керування.

Оскільки процесам керування інформації притаманні властивості випадковості, то при вивченні цих процесів широко використовуються ймовірнісні методи, причому не лише класичні, а й виникає потреба у створенні нових. Отже, теорія інформації може розглядатись з одного боку як окрема прикладна наука, яка використовує ймовірнісні методи, а з іншого – як розділ теорій ймовірностей.

На нашу думку, при навчанні курсу «Теорія ймовірностей, математична статистика та випадкові процеси» студентів спеціальності „Програмування електронно-обчислювальної техніки та автоматизованих систем" (ПЗАС) доцільним є ознайомлення студентів з основними поняттями теорії інформації, що сприятиме формуванню математичної та інформаційної культури, підвищуватиме мотивацію навчання, створить відповідну теоретичну базу для вивчення ряду професійно-орієнтованих дисциплін.

Курс «Теорія ймовірностей, математична статистика та випадкові процеси» є фундаментальним курсом при підготовці студентів освітньо-кваліфікаційного рівня „бакалавр" спеціальності ПЗАС. Він включає наступні розділи: випадкові події, ймовірність випадкової події; випадкові величини; основні закони розподілу; граничні теореми; елементи математичної статистики; елементи дисперсійного, кореляційного та регресійного аналізу; елементи теорії випадкових процесів.

## 1. Вивчення елементів теорії інформації (ймовірнісний підхід)

Однією з основних задач теорії інформації є відшукання найбільш економних методів кодування, що дозволяють передати задану інформацію за допомогою мінімальної кількості символів. Інша типова задача має наступне формулювання: в наявності є джерело інформації (передавач), який безперервно виробляє інформацію, і канал зв'язку, по якому ця інформація передається в іншу інстанцію (приймач). Якою має бути пропускна здатність каналу зв'язку, щоб канал „справлявся" зі своєю задачею?

Ці задачі можуть бути використані як проблемні питання для обґрунтування необхідності подальшого розгляду нових імовірнісних категорій, або як ілюстрація застосування імовірнісних методів до прикладних задач.

Введення елементів теорії інформації у курсі теорії ймовірностей можна здійснити на основі розкриття наступних питань:

1. Основні задачі теорії інформації.
2. Поняття ентропії та її властивості.
3. Ентропія складної системи. Теорема додавання ентропій.
4. Умовна ентропія.
5. Ентропія і інформація.
6. Задачі кодування повідомлень. Код Шеннона-Фено.

Щоб бути готовими до сприйняття нового матеріалу у студентів мають бути вже сформовані наступні поняття: ймовірність випадкової події, теорема множення ймовірностей незалежних подій, випадкові величини (неперервні, дискретні), числові характеристики випадкових величин, зокрема, математичне сподівання, двовимірні випадкові величини, умовні ймовірності.

Вивчення елементів теорії інформації можна розпочати з постановки задачі кодування повідомлень, реалізовуючи проблемний підхід у навчанні цього розділу. При цьому ввести поняття коду та кодування, елементарних символів, системи Х, яка кодується (наприклад букви алфавіту), системи У, за допомогою якої кодується (розглянути введені поняття на прикладі азбуки Морзе). Міркуючи логічно студенти самостійно можуть зробити висновок щодо співвідношення кількості можливих станів системи Х та У.

З циклу професійно-орієнтованих та математичних дисциплін, що вивчались раніше, студенти вже ознайомлені з різними формами передавання сигналів, з різними системами числення, зокрема, двійковою та десятковою, вміють переводити числа з однієї системи числення в іншу.

На цьому етапі необхідно запропонувати студентам наступну задачу: закодувати двійковим кодом літери алфавіту, так щоб кожній літері відповідала певна комбінація елементарних символів 0 та 1 і щоб середнє число цих символів на літеру тексту було мінімальним. І поставити додатково наступне завдання: після побудови коду дослідити чи є запропонований код дійсно оптимальним?

Для першого завдання студентам потрібні лише логічні міркування та знання двійкової системи числення.

Усім літерам алфавіту пропонується приписати номери від 0 до 31 (для російського алфавіту), плюс проміжок між словами „-” під номером 32. Далі перевести нумерацію з десяткової системи числення у двійкову.

В такому коді на зображення кожної букви відводиться рівно 5 символів.

Поставити проблемне запитання: чи не існує іншого коду, і чи є запропонований код оптимальним? Також для наштовхування на правильну відповідь запитати чи всі літери однаково часто зустрічаються у словах. Відповідь на дане питання має дати змогу студентам сформулювати гіпотезу, що можливо існує інший більш оптимальний код, який враховує частоти появ тих чи інших літер, і для літер, що зустрічаються частіше, можна було б знайти код, що витрачає менше число елементарних символів, а для літер, що зустрічаються рідше – більше число символів коду. Після чого запропонувати студентам таблицю частот літер.

Очевидно такий код буде економнішим, але щоб це обґрунтувати необхідно володіти таким поняття як інформація.

Введення поняття інформації базується на понятті ентропії як міри невизначеності стану деякої фізичної системи.

Розглянемо деяку систему Х, яка може приймати скінченну множину станів: $x_1, x_2, ..., x_n$ з ймовірностями $p_1, p_2, ..., p_n$, де $p_i = P(X \sim x_i)$ - ймовірність того, що система Х прийме стан $x_i$, Х~$x_i$ - подія {система знаходиться в стані $x_i$}. Очевидно $\sum_{i=1}^{n} p_i = 1$.

Запишемо ці данні у вигляді таблиці :

| | | | | |
|---|---|---|---|---|
| $x_i$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ |
| $p_i$ | $p_1$ | $p_2$ | $\cdots$ | $p_n$ |

Такий запис зрозумілий і вже знайомий студентам, оскільки є аналогічним до зображення ряду розподілу дискретної випадкової величини Х. Відмінним є лише те, що для визначеності ступеня невизначеності системи зовсім не важливо самі значення, важлива лише їх кількість.

Ентропією системи називається сума добутків ймовірностей різних станів системи на логарифми цих ймовірностей, взята з протилежним знаком:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i$$

Логарифм у формулі можна брати за довільною основою $a$ >1.Якщо за основу обрано число 10, то говорять про «десяткові одиниці» ентропії. На практиці, як правило, використовують логарифми за основою 2 тоді вимірюється ентропія в «двійкових одиницях».

Вибір саме логарифмічної функції можна пояснити наступними міркуваннями. Розглянемо випробування, яке має $k$ рівноможливих результатів. Зрозуміло, що коли $k=1$, результат випробувань не є випадковим і жодної невизначеності немає. Зі збільшенням $k$ невизначеність зростає. Отже, числова характеристика невизначеності $f(k)$ має бути $f(1)=0$ 0 і зростати зі збільшенням $k$. Розглянемо два незалежні випробування α і β. Нехай випробування α має *m*, а випробування β — *n* результатів. Добуток подій αβ матиме *mn* результатів. Невизначеність випробування αβ буде більшою і від α, і від β. Природно припустити, що ступінь невизначеності випробування αβ дорівнює сумі невизначеностей, які характеризують випробування α і β. Звідси дістаємо таку умову: $f(mn)=f(m)+f(n)$. Найпростішою функцією, що має вказану властивість є логарифмічна функція.

**Приклад**. Визначити ентропію фізичної системи, що складається з двох стрільців, які стріляють по мішені. В результаті змагання система може опинитись в одному з двох можливих станів:

1. обоє стрільців влучили;
2. перший влучив, другий не влучив;
3. другий влучив, перший не влучив;
4. обидва не влучили.

Ймовірності цих станів відповідно дорівнюють 0,2; 0,3; 0,4; та 0,1.

*Розв'язання:*

| $x_i$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $p_i$ | 0,2 | 0,3 | 0,4 | 0,1 |

За формулою ентропії маємо:

$$H(X)=-p_1\log p_1-p_2\log p_2-p_3\log p_3-p_4\log p_4$$

$$H(X)=-0{,}2\log 0{,}2-0{,}3\log 0{,}3-0{,}4\log 0{,}4-0{,}1\log 0{,}1$$

$$H(X)=0{,}4644+0{,}5211+0{,}5288+0{,}322\approx 1{,}85\,(\text{дв. од.})$$

Після введення означення необхідно обґрунтувати наступні властивості ентропії:

- ентропія обертається в нуль, коли один зі станів системи достовірний, а інші – неможливі;
- при заданій кількості станів ентропія обертається в максимум, якщо ці стани рівноймовірні;
- при збільшенні кількості станів значення ентропії зростає;
- ентропія також володіє властивістю адитивності, тобто якщо декілька незалежних систем об'єднати в одну, їх ентропії додаються.

Ознайомившись з поняттям ентропії можна ввести поняття інформації та її кількості.

Перед введенням поняття «кількості інформації» студенти вже ознайомлені з поняття ентропії як міри невизначеності стану деякої фізичної системи. Очевидним є припущення, що кількість інформації вимірюється зменшенням ентропії тієї системи, для уточнення стану якої призначені відомості. Оскільки, при отриманні відомостей невизначеність системи може бути зменшена, й чим більший об'єм отриманих відомостей, чим вони є змістовнішими, тим більше інформації, тим менша невизначеність системи.

Розглянемо систему X, до отримання відомостей про яку ентропія системи становила $H(X)$. Після отримання відомостей стан системи повністю визначився, тобто ентропія стала рівна нулю. Нехай $I_x$ - інформація, що отримується в результаті з'ясування станів системи X. Вона дорівнює зменшенню ентропії, тобто *кількість інформації, яку ми набуваємо при повному з'ясуванні стану деякої фізичної системи, дорівнює ентропії цієї системи.*

$$I_x = H(X)$$

$$I_x = -\sum_{i=1}^{n} p_i \log p_i$$

Інформація $I_x$ є усередненим за всіма станами системи логарифму ймовірностей стану з протилежним знаком.

Кожен доданок $-\log p_i$ розглядається як частинна інформація, що отримується від окремого повідомлення, яке полягає у тому, що система X знаходиться у стані $x_i$. Позначимо таку інформацію $I_{x_i}$:

$$I_{x_i} = -\log p_i$$

Тоді інформація $I_x$ подається як середня (або повна) інформація, що отримується від всіх можливих окремих повідомлень з врахуванням їх ймовірностей.

Колмогоров А.Н. розглядав три підходи до введення поняття інформації: комбінаторний; ймовірнісний; алгоритмічний.

Комбінаторний підхід має певну логічну незалежність від будь-яких ймовірнісних припущень, проте надання змінним характеру «випадкових змінних», що володіють певним розподілом ймовірностей, надає можливість отримати більш багатшу систему понять та співвідношень. Тому при викладенні елементів теорії інформації поняття «кількості інформації» краще вводити, дотримуючись ймовірнісного підходу.

Проте у ймовірнісному підході має місце один парадокс – при комбінаторному підході величина $I_x$ завжди невід'ємна, що є природним при звичайному уявленні про «кількість інформації», при ймовірнісному підході ця величина може бути і від'ємна. Справжньою мірою «кількості інформації» є усереднена величина $I_x$.

**Приклад**. Визначити частинну інформацію, що міститься у повідомленні особи А, яку вперше зустріли: «сьогодні мій день народження».

*Розв'язання.* Всі дні у році можуть бути з однаковими ймовірностями днями народження особи А. Ймовірність отриманого повідомлення $p = \frac{1}{365}$.

Частинна інформація від даного повідомлення $i = -\log \frac{1}{365} \approx 8{,}51$

Оскільки на практиці часто доводиться визначати ентропію для складної системи , що отримана шляхом об'єднання двох або більше простих систем, то наступною темою розгляду понять теорії інформації є «Ентропія складної системи. Умовна ентропія».

Зображення об'єднання двох систем з їх ймовірностями перебування в певному стані є аналогічними до двовимірних випадкових величин, що вивчались студентами в основному курсі теорії ймовірностей. Визначення ентропії для таких систем має певну аналогію з обчисленням числових характеристик умовних законів розподілу.

### 2. Альтернативні підходи до визначення кількості інформації

Ймовірнісна (статистична) теорія інформації пов'язує поняття інформації зі зниженням невизначеності (ентропії) стану об'єкта. Підходи й математичний апарат для кількісного визначення інформації та ентропії, що їх розробили К. Шенон та Н. Вінер, виявилися корисними в технічних застосуваннях (теорії зв'язку) — оптимізації кодування, передавання, зберігання інформації тощо. Їх праці з теорії інформації сприяли розумінню того, що не існує абсолютної інформації про об'єкт, визначення інформації залежить від вибраної моделі об'єкта. Оскільки залежно від мети дослідження вибирають різні моделі з різним описом станів об'єкта, то й з'ясування інформації про об'єкт залежить від мети та завдань дослідника. Адже в одних і тих самих даних міститься різна кількість інформації для різних завдань управління. Однак статистична теорія інформації не набула поширення для задач обробки інформації, призначеної для управління економічними об'єктами. Це пояснюється тим, що її підходи не враховують специфіки економічної інформації (зокрема, відкидаються змістовні взаємозв'язки, ігнорується зміст та корисність інформації для досягнення мети — цінність, доцільність). Наприклад, кількість інформації на символ є лише усередненою мірою невизначеності появи цього символа. Тому загальна кількість інформації, що міститься в деякому повідомленні ($I = -\Sigma p_i \log p_i$), ніяк не пов'язується зі змістовністю і корисністю цієї інформації для одержувача. Інформативність повідомлень для одержувача залежить від його сфери інтересів, роду занять, мети дослідження тощо. Отже, необхідно враховувати різні аспекти оцінки кількості інформації: не лише за формально-структурними ознаками, а й за змістом та практичною цінністю для одержувача.

Однією з найбільш важливих властивостей інформації є її корисність. Але бути корисним може тільки те, що має сенс для даної системи. Реальні (зокрема, економічні) системи перебувають у процесі постійного перетворення, причому будь-яке елементарне перетворення в системі є подією. Кожна подія супроводжується повідомленням, яке є інформаційним еквівалентом події. З огляду на сказане інформація — це повідомлення, яке має сенс для даної системи. Але значення будуть мати тільки ті повідомлення, які обмежують різноманітність поводження досліджуваної системи в напрямку її пристосування до середовища.

Наведемо деякі інші міркування та підходи до визначення кількості інформації.

Семантичний підхід. Один із методів обчислення кількості семантичної інформації полягає в тому, щоб визначати її через так звану логічну імовірність, що являє собою ступінь підтвердження тієї чи іншої гіпотези. При цьому кількість семантичної інформації, що міститься в повідомленні, зростає зі зменшенням ступеня підтвердження гіпотези. Отже, якщо логічна ймовірність дорівнює одиниці, тобто якщо вся гіпотеза побудована на відомих даних та цілком підтверджується повідомленням, то таке повідомлення не приносить адресатові нічого нового і семантична інформація дорівнює нулю. (Наприклад, повідомлення «Дніпро впадає в Чорне море».) І навпаки, зі зменшенням ступеня підтвердження гіпотези (чи, інакше кажучи, апріорного знання) кількість семантичної інформації, що її доставляє повідомлення, зростає.

З описаним підходом до визначення інформаційної змістовності повідомлень стикається запропонована Ю. Шрейдером ідея, що ґрунтується на врахуванні «запису знань» (тезауруса) одержувача. Тезаурусом (грец. «скарб») називають словник, в якому наведено не тільки значення окремих слів, а й змістовні зв'язки між ними (наприклад, тлумачний словник Даля). У розглядуваному контексті під тезаурусом розуміють деякий узагальнений довідник, що визначає рівень знань одержувача про повідомлення. При цьому повідомлення, що містять нову для одержувача інформацію, змінюють, збагачують його тезаурус. Якщо повідомлення не вносить нічого нового в тезаурус одержувача, то природно вважати, що змістовна семантична інформація дорівнює нулю. Якщо одне з двох повідомлень змінює тезаурус незначно, а друге вносить до нього істотні зміни, то природно вважати, що друге повідомлення є змістовнішим, несе в собі значно більший обсяг семантичної інформації. При цьому під зміною тезауруса слід розуміти не тільки появу нових понять, а й встановлення нових зв'язків між ними, ліквідацію застарілих понять чи зв'язків тощо.

Прагматичний підхід. Визначаючи інформацію, ми зазначали, що однією з властивостей інформації є використання її у процесах управління. А коли інформація використовується для управління, то її, природно, належить оцінювати з позицій корисності, цінності, доцільності для досягнення поставленої мети управління. Тому кожне одержуване ланками управління повідомлення важливо оцінювати не з погляду пізнавальних характеристик, а з прагматичного, тобто з боку корисності чи цінності для виконання функцій управління.

Виходячи з таких міркувань, А. Харкевич запропонував міру цінності інформації Іц визначати як зміну ймовірності досягнення мети в разі отримання цієї інформації:

$$I_{ц} = \log p_1 - \log p_0 = \log \frac{p_1}{p_0},$$

де $p_0$ — початкова (до отримання відомостей) імовірність досягнення мети;

$p_1$ — імовірність досягнення мети після отримання інформації.

При цьому можливі три різні випадки:

1. Отримана інформація не змінює ймовірності, тобто $p_1 = p_0 \Rightarrow$ $I$ц = 0. Таку інформацію називають порожньою.
2. Якщо імовірність досягнення мети збільшується: $p_1 > p_0 \Rightarrow$ $\Rightarrow$ $I$ц > 0, то прагматична інформація зросла.

3. Якщо ймовірність зменшилася: $p_1 < p_0 \;\Rightarrow\; I\text{ц} < 0$, це означає, що отримана інформація є негативною, тобто дезінформацією.

Зауважимо, що прагматичні та семантичні оцінки важко розмежувати, а в деяких випадках вони збігаються.

## Висновки

Вивчення елементів теорії інформації в курсі «Теорія ймовірностей, математична статистика та випадкові процеси» студентів спеціальності „Програмування електронно-обчислювальної техніки та автоматизованих систем” є природним і доцільним, це дозволить:

- підвищити рівень математичної та інформаційної культури студентів;
- покращити мотивацію навчання;
- активізувати пізнавальну діяльність студентів;
- реалізувати міжпредметні зв’язки комп’ютерно-орієнтовних та математичних дисциплін;
- розв’язати ряд прикладних та професійно-орієнтованих задач;
- урізноманітнити форми проведення занять (можна ввести ряд лабораторних робіт з написанням програми й математичним обґрунтуванням);
- ефективніше організувати самостійну роботу студентів.

## Список використаної літератури

1. Гнеденко Б.В. Курс теории вероятности. — М.: Физматгиз, 1961.
2. Яглом А.М., Яглом И.М. Вероятность и информация. — М.: Гостехиздат, 1960.
3. Хэмминг Р.В. Теория кодирования и теория информации. — М.: Радио и связь, 1983.
4. Шеннон К. Работы по теории информации и кибернетике. — М.: Изд-во иностранной литературы, 1963.
5. Берлекэмп Э. Алгебраическая теория кодирования. М.: Мир, 1971.
6. Бриллюэн Л. Наука и теория информации. — М.: Физ.-мат.лит., 1980.
7. Галлагер Р. Теория информации и надежная связь. — М.: Советское радио, 1974.
8. Гоппа В.Д. Введение в алгебраическую теорию информации. — М.: Наука, 1995.
9. Колмогоров А.Н. Теория информации и теория алгоритмов. — М.: Наука, 1987.