ЛЕКЦІЯ 2.
Аксіоматика теорії ймовірностей

1. *Геометричне визначення ймовірності.*

В *лекції* 1, в третьому серед прикладів стохастичних експериментів розглядається ситуація, пов'язана з *моментом надходження* до системи обслуговування першої вимоги. Простором елементарних подій тоді буде *незлічена множина* $\Omega = [0, +\infty)$ всіх невід'ємних чисел.

Тому в цьому випадку запропоновану процедуру визначення ймовірностей на *злічених просторах* елементарних наслідків неможливо реалізувати:

- По-перше – елементарних подій «*занадто*» багато, щоб кожній приписати порядковий номер;
- А по-друге – кожній окремій такій події $\omega \in \Omega$ необхідно було б визначити нульову ймовірність:

$$P(\omega) = 0,\ \omega \in \Omega.$$

Однак існує велика кількість окремих стохастичних експериментів, що мають *незлічені простори* елементарних наслідків Ω, для яких можна:

- Конструктивним чином визначити множину $\mathfrak{I}$ підмножин простору Ω:

$$\mathfrak{I} = \{A \,/\, A \subseteq \Omega\},$$

  елементи $A \in \mathfrak{I}$ якої будемо вважати *випадковими подіями*:
- Запропонувати відповідну процедуру знаходження числового значення ймовірності $P(A)$ для кожної випадкової події $A \in \mathfrak{I}$.

один такий, що

В якості прикладу розглянемо схему, що дістала назву

➢ *геометричного визначення ймовірності.*

Це визначення можна розглядати як аналог *класичного визначення* у випадку *нескінченої* кількості можливих наслідків. При цьому така «*аналогія*» повинна розглядатися в наступному розумінні:

- Якщо формулювання «*випадковим чином*» для скінчених просторів елементарних подій означає на практиці, що «*всі наслідки так само ймовірні*» в сенсі *класичного визначення ймовірності*,
- То, як правило, термін «*випадковим чином вибираємо точку в* $\Omega \subset R^n$» означає, що «*всі наслідки такого відбору так само ймовірні*» в сенсі *геометричного визначення ймовірності.*

Приведемо точні формулювання.

- Припустимо, що простір елементарних наслідків $\Omega \subset R^n$ є обмеженою множиною *n–вимірного Евклідового простору*.

Іншими словами, якщо $m(A)$ позначає міру цієї множини (на прямій ($n = 1$) – це довжина відрізка, на площині ($n = 2$) – це площа фігури, в просторі ($n = 3$) – це об'єм області, ітп.), то:

$$m(\Omega) < \infty.$$

- Припустимо також, що всі елементарні наслідки $\omega \in \Omega$ *рівно-можливі*.

Зауважимо, що вираз «*рівно-можливі*», або «*так само правдоподібні*», в цьому випадку розуміємо наступним чином:

  - *немає підстав вважати одні серед них більш ймовірними в порівнянню з іншими*.

- *Випадковими подіями* будемо вважати всі підмножин $A \subseteq \Omega$, для яких міра $m(A)$ визначена.

Тоді згідно з геометричним визначенням ймовірність $P(A)$ будь-якої випадкової події A визначається формулою:

$$P(A) = \frac{m(A)}{m(\Omega)}, A \subset \Omega.$$

Звернемо увагу, що визначену таким чином ймовірність, так само як і в попередніх випадках, характеризують наступні властивості:

- $P(A) \geq 0$ для довільної події $A \subseteq \Omega$.
- $P(\Omega) = 1$.
- Якщо події A та B – несумісні, то $P(A \cup B) = P(A) + P(B)$.

2. *Приклади використання геометричного визначення.*

Наведемо декілька прикладів практичного використання геометричного визначення ймовірності.

Приклад 1. (***Задача про зустріч***). Припустимо, що дві особи $\mathbf{O_1}$ і $\mathbf{O_2}$ домовилися зустрітися в проміжку часу $[0, T]$. Той, хто приходить першим, чекає зустрічі протягом τ одиниць часу, а потім залишає місце зустрічі.

Припускаючи, що моменти прибуття осіб на зустріч є незалежними точками, випадково вибраними в інтервалі $[0, T]$, слід знайти ймовірність того, що: *зустріч відбудеться*.

Розв'язок. Нехай x означає моментом прибуття на зустріч особи $\mathbf{O_1}$, а y – відповідно момент прибуття особи $\mathbf{O_2}$.

Тоді в якості простору Ω елементарних наслідків стохастичного експерименту, що розглядається, можна вибрати квадрат з вершинами в точках:

$$(0, 0), (0, T), (T, T), (T, 0)$$

прямокутної системи координат XOY:

$$\Omega = \{(x, y): 0 \leq x \leq T; 0 \leq y \leq T\}.$$

Мірою простору Ω буде площа квадрату:

$$m(\Omega) = T^2 < \infty.$$

Позначимо літерою D подію, яка нас цікавить:

$$D = \{\textit{Запланована зустріч відбудеться}\}.$$

Тоді формально підмножина елементарних наслідків, що сприяють випадковій події D, записується наступним чином:

$$D = \{(x, y) \in \Omega: |x - y| \le \tau\}.$$

Легко переконатися, що площа цієї множини дорівнює:

$$m(D) = T^2 - (T - \tau)^2.$$

Отже згідно з геометричним визначенням ймовірності отримаємо:

$$P(D) = \frac{m(D)}{m(\Omega)} = \frac{T^2 - (T-\tau)^2}{T^2} = 1 - \left(1 - \frac{\tau}{T}\right)^2.$$

Приклад 2. (***Задача Бюффона***). Припустимо, що площину покривають паралельні прямі, що проходять на відстані $2a$ один від одної.

Голку довжиною $2l$, ($l < a$) кидаємо випадковим чином на цю площину.

Знайти ймовірність того, що: *голка перетне одну з прямих*.

Розв’язок. Нехай x означає відстань від центру голки до найближчої прямої, а φ – кут між голкою та напрямком прямих.

Тоді пара чисел (x,φ) може бути елементарним наслідком стохастичного експерименту, пов’язаного з киданням голки:

- ➢ *Ці числа визначають положення голки на площині і дають можливість встановити, відбулась досліджувана випадкова подія, чи ні.*

Виходячи з припущень, відстань (x) від центру голки від найближчої прямої не перевищує a, а кут (φ) може змінюватися від 0 до 180°. Тому простором елементарних подій (Ω) розглянутого стохастичного експерименту буде наступний прямокутник:

$$\Omega = \{(x, \varphi): 0 \le x \le a; 0 \le \varphi \le \pi\}.$$

Припущення:

- «*кидаємо голку випадковим чином*»

означає, що всі точки $\omega = (x, \varphi) \in \Omega$ є «*однаково ймовірними*», тобто:

- *не маємо підстав надати перевагу будь-яким серед них.*

Тому можна скористатися з геометричного визначення ймовірності, відповідно до якого мірою множини Ω буде її площа:

$$m(\Omega) = a \cdot \pi < \infty.$$

Позначимо літерою C подію, яка нас цікавить:

$$C = \{\textit{Голка перетне одну з прямих}\},$$

та знайдемо підмножину елементарних наслідків $\omega = (x, \varphi) \in C$, що сприяють випадковій події C.

Враховуючи умову ($l < a$), приходимо до висновку, що голка може перетнути тільки одну («*найближчу*» до неї) пряму.

Відстань від центру голки до цієї прямої визначає координата x елементарного наслідку $\omega = (x, \varphi) \in C$. Кут φ визначає, чи перетне її голка, чи ні:

- *Подія C відбудеться тоді і тільки тоді, коли*: $x \le l \cdot \mathrm{Sin}(\varphi)$.

Підсумовуючи можемо формально подію C записати таким чином:

$$C = \{\textit{Голка перетне одну з прямих}\} = \{(x, \varphi) \in \Omega\colon x \le l \cdot \mathrm{Sin}(\varphi)\}.$$

Мірою $m(C)$ множини C буде її площа:

$$m(C) = \int_0^{\pi} l \cdot Sin(\varphi) d\varphi = 2l.$$

Використовуючи, подібно як і в попередніх прикладах, геометричне визначення ймовірності, отримаємо:

$$P(C) = \frac{m(C)}{m(\Omega)} = \frac{2l}{a\pi}.$$

Варто зауважити, що описаний стохастичний експеримент з «*киданням голки*» дає можливість *експериментальним шляхом* знайти наближення для значення числа π:

- Для цього необхідно в незмінних умовах повторити «*достатньо багато разів*» цей експеримент та замінити в отриманій формулі ймовірність $P(C)$ її статистичною оцінкою.

3. *Геометричні парадокси*.

Якщо навіть у випадку класичного визначення ймовірності, коли кількість елементарних наслідків скінчена, перевірка їх *рівнозначності* не завжди є простим завданням, то, тим більше, у випадку незлічених множин подібна перевірка часто може бути дуже складна. Причиною цього є та обставина, що дуже складно *однозначно* описати формальними засобами, що означає:

- «*випадковий*» *вибір серед незліченої кількості елементів*.

Ця обставина була приводом до великої кількості «*парадоксів*». Найбільш відомим з цієї точки зору можна вважати «*парадокс Бертрана*».

Виявляється, що остаточна відповідь суттєво залежить, від згаданого визначення того, що слід вважати:

- «*випадковим вибором серед незліченої кількості елементів*».

Приклад 3. «***Парадокс Бертрана***».

- Маємо задане коло.
- «*Випадковим чином*» вибираємо в цьому колі хорду.

Наскільки ймовірно, що:

- Довжина вибраної *хорди* буде *більша* від довжини *сторони* вписаного в це коло *рівностороннього трикутника*.

Приведемо кілька цілком обґрунтованих розв’язків, які, однак, дають абсолютно різні відповіді.

Варіант 1. З міркувань *симетрії* можна спочатку вибрати конкретний напрямок хорди. Зробимо це.

Після цього проведемо *діаметр*, перпендикулярний до вибраного напрямку. Очевидно, що при такому розумінні

➢ «*випадкового вибору серед незліченої кількості елементів*.

сприятливими події:

«*Довжина хорди більша від довжини сторони трикутника*»

будуть хорди:

що проходять в інтервалі ***від ¼ до ¾*** *збудованого діаметру*.

Тобто рівно половина всіх можливих наслідків. Отже при такому розумінні «*випадкового вибору*» **ймовірність** того, що довжина хорди буде *більша* від довжини вписаного в це коло рівностороннього трикутника:

дорівнює ½.

Варіант 2. З міркувань *симетрії* можна спочатку «*закріпити*» на колі один кінець хорди. Зробимо це.

Після цього проведемо *дотичну до кола* в цій точці, та впишемо в коло *рівносторонній трикутник* з вершиною в цій точці. Дотична буде утворювати зі сторонами трикутника кути по $60°$.

Очевидно, що при такому розумінні «*випадкового вибору*» сприятливими для нашої події будуть хорди, що потрапили до *середнього* серед цих кутів.

Отже **ймовірність** того, що довжина хорди буде *більша* від довжини вписаного в це коло рівностороннього трикутника:

дорівнює $^1/_3$.

Варіант 3. Розташування хорди можна однозначно визначити, вибираючи її середину. Множину сприятливих для нашої події наслідків тоді будуємо наступним чином.

- Проведемо концентричне коло, що проходить через *середину радіуса*.
- Довжина хорди буде *більша* від довжини вписаного в це коло рівностороннього трикутника, якщо її центр потрапить до *меншого* кола.

Площа *меншого* кола становить ¼ площі даного кола.

Отже **ймовірність** того, що довжина хорди буде *більша* від довжини вписаного в це коло рівностороннього трикутника:

дорівнює ¼.

4. *Аксіоми теорії ймовірностей.*

Залежність способу визначення ймовірності від того явища, яке досліджується, дуже обмежене коло застосувань наведених вище можливих підходів до визначення поняття:

«*ймовірність випадкової події*»,

відсутність *точних математичних формулювань* і як наслідок – величезна кількість протиріч та парадоксів, спонукали математиків до пошуків більш плідних розв'язків сформульованої вище *проблеми формалізації ймовірності*.

Вимагав цього швидкий розвиток нових галузей науки та техніки, зокрема успіхи в ядерній фізиці, необхідність досліджень соціально-економічних та демографічних явищ. Характерною рисою цих явищ були умови *невизначеності* та *непередбачуваності* їх наслідків.

Тому невипадково задача *формалізації основ теорії ймовірності* потрапила до славнозвісного *переліку великих проблем Гілберта*.

- Необхідно було чітко і точно визначити всі пов'язані із стохастичними експериментами поняття;
- Вибудувати їх логічні взаємозв'язки;
- Та з'ясувати умови, що гарантують можливість прийняття обґрунтованих рішень з допомогою створеної в результаті цього теорії.

Єдине що не викликало сумнівів, так це чітке усвідомлення наступної обставини:

  - Несуперечливу, плідну *теорію ймовірностей* неможливо створити, відштовхуючись від *якогось конкретного одиничного стохастичного експерименту*.

Необхідно було будувати *абстрактну дисципліну*, визначення базових понять якої *не залежить* від можливих властивостей тих чи інших випадкових явищ.

Найбільш плідним з цієї точки зору виявився, в певному розумінні «*традиційний*» в математиці, *аксіоматичний підхід*. Остаточний розв'язок належить радянському математику А.М. Колмогорову, який запропонував в 1933 р. здавалося б дуже *просту* та *природню*, а водночас дуже *ефективну* та *плідну*, систему аксіом теорії ймовірностей.

Має вона наступний вигляд.

- Базовими в цій *аксіоматиці* є поняття *стохастичного експерименту*, *випадкової події* та *ймовірності випадкової події*.
  - Припустимо, що задано деякий простір елементарних наслідків Ω, тобто множину $\Omega = \{\omega\}$, елементами якої є елементарні події ω.
  - Припустимо далі, що визначена певна множина $\mathfrak{I}$ підмножин простору Ω: $\mathfrak{I} = \{A \,/\, A \subseteq \Omega\}$, елементи $A \in \mathfrak{I}$ якої називаємо *випадковими подіями* і для якої виконуються наступні умови:

A1. $\Omega \in \mathfrak{I}$.

A2. Якщо $A \in \mathfrak{I}$, то $\overline{A} = \Omega \setminus A \in \mathfrak{I}$.

A3. Якщо $A_1 \in \mathfrak{I}$, $A_2 \in \mathfrak{I}$, $A_3 \in \mathfrak{I}$, ...то:

$$\bigcup_{i=1}^{\infty} A_i \in \mathfrak{I}.$$

Сукупність множин, для якого виконані умови **A1**, **A2**, **A3**, називається *σ-алгеброю* множин.

- Припустимо, далі, що на множині $\mathfrak{I}$ визначена деяка дійсна функція:

$$P: \mathfrak{I} \to [0, 1],$$

для якої виконуються наступні умови:

P1. $P(A) \geq 0$ для $A \in \mathfrak{I}$.

P2. $P(\Omega) = 1$.

P3. Якщо $A_1 \in \mathfrak{I}$, $A_2 \in \mathfrak{I}$, $A_3 \in \mathfrak{I}$, ... – попарно несумісні випадкові події, тобто $A_i \cap A_j = \varnothing$, якщо $i \neq j$, то:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

- Значення $P(A)$ функції P для $A \in \mathfrak{I}$ називаємо *ймовірністю випадкової події A*.
- Трійку $(\Omega, \mathfrak{I}, P)$, де Ω, це простір елементарних наслідків, $\mathfrak{I}$, це множина *випадкових подій*, а P, це *ймовірність*, називаємо *ймовірнісним простором*.
- Умови **A1**, **A2**, **A3**, **P1**, **P2**, **P3**, утворюють *систему аксіом* теорії ймовірностей.

5. *Властивості ймовірності.*

Легко переконатись, що для всіх трьох визначень ймовірності: *класичного*, *на злічених просторах* Ω та *геометричного*, виконуються всі перераховані аксіоми. Це є підтвердженням того, що:

- Існують *реальні утворення*, які задовольняють всім абстрактним вимогам *аксіоматичного визначення ймовірності*.

Отже маємо приклади *конкретних ймовірнісних просторів*.

Сформульовані аксіоми – це *найпростіші*, *очевидні* властивості ймовірності, що не потребують доведення.

Спираючись на них можна доводити все нові і нові, більш складні і далеко не очевидні властивості ймовірності. Таким чином маємо змогу швидко розбудовувати *змістовну* математичну теорію. Приведемо кілька прикладів, що ілюструють це.

Властивість 1. Якщо ймовірність випадкової події A дорівнює $P(A)$, то ймовірність протилежної події $\overline{A} = \Omega \setminus A$ обчислюємо за формулою:

$$P(\overline{A}) = P(\Omega \setminus A) = 1 - P(A).$$

Доведення. Оскільки $A \cup \overline{A} = \Omega$, і $A \cap \overline{A} = \varnothing$, то:

$$1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A}).$$

Властивість 2. Ймовірність неможливої події дорівнює нулю: $P(\varnothing) = 0$.
Доведення. Оскільки $\overline{\Omega} = \varnothing$, то:

$$P(\varnothing) = P\left(\overline{\Omega}\right) = 1 - P(\Omega) = 1 - 1 = 0.$$

Властивість 3. Якщо $A \subset B$, то $P(B\backslash A) = \mathrm{P}(B) - \mathrm{P}(A)$.
Доведення. З визначення різниці $(B\backslash A)$ подій A та B отримаємо:

$$B = (B\backslash A) \cup A,\ (B\backslash A) \cap A = \varnothing.$$

Таким чином на підставі аксіоми **P3**: $P(B) = P(B\backslash A) + P(A)$.

Оскільки згідно аксіом ймовірність $P(B\backslash A) \geq 0$ невід'ємна, то з останньої рівності відразу отримаємо наступну властивість.
Властивість 4. Якщо $A \subset B$, то $P(B) \geq P(A)$.

Якщо замінити в останній нерівності B на Ω, то на підставі аксіоми **P2** відразу отримаємо наступну властивість.
Властивість 5. Для кожної події $A \in \mathfrak{I}$:

$$P(A) \leq 1.$$

Властивість 6. (*Формула додавання ймовірностей*). Для довільних подій $A \in \mathfrak{I}$ та $B \in \mathfrak{I}$ справедлива рівність:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Доведення. Безпосередньо з визначення різниці подій A та B отримаємо наступну рівність:

$$A \cup B = \{A\backslash(A \cap B)\} \cup (A \cap B) \cup \{B\backslash(A \cap B)\}.$$

Випадкові події $A\backslash(A \cap B)$, $A \cap B$ та $B\backslash(A \cap B)$ попарно несумісні, крім того $A \cap B \subset A$ та $A \cap B \subset B$. Отже:

$$P(A \cup B) = P(A\backslash(A \cap B)) + P(A \cap B) + P(B\backslash(A \cap B)) =$$
$$= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) =$$
$$= P(A) + P(B) - P(A \cap B).$$

Використовуючи метод математичної індукції легко переконатися, що у випадку, коли кількість «*доданків*» більша ніж 2, то «*формула додавання ймовірностей*» приймає наступний вигляд.
Властивість 7. (*Формула додавання ймовірностей*). Нехай $A_1, A_2, ..., A_n \in \mathfrak{I}$ – довільні випадкові подій. Тоді має місце наступна рівність:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<k}^{n} P(A_i \cap A_k) +$$
$$+ \sum_{i<k<j}^{n} P(A_i \cap A_k \cap A_j) - \ldots (-1)^{n-1} \cdot P(A_1 \cap A_2 \cap \ldots \cap A_n)\,).$$

Властивість 8. (*Неперервність ймовірностей*). Припустимо, що маємо послідовність випадкових подій A_i, $i = 1, 2, ..., n, ...$, $(A_1, A_2, ..., A_n, ... \in \mathfrak{I})$, для яких виконується наступна умова:

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots.$$

Тоді має місце наступна рівність:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n).$$

Доведення. Введемо послідовність випадкових подій $B_1, B_2, \ldots, B_n, \ldots \in \Im$ наступним чином:

$$B_1 = A_1,\ B_2 = A_2 - A_1,\ B_3 = A_3 - A_2,\ \ldots\ B_n = A_n - A_{n-1},\ \ldots$$

Приймаючи до уваги «*зростаючий*» характер послідовності випадкових подій A_i, $i = 1, 2, \ldots, n, \ldots$, отримуємо наступні рівності:

$$\bigcup_{i=1}^{n} B_i = A_n \text{; та } \bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i .$$

Крім того, очевидно, що $B_k \cap B_l = \varnothing$, якщо $k \neq l$ Тому на підставі аксіоми (**P3**) маємо:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(A_n).$$

Застосовуючи «*правила де-Моргана*», отримаємо наступний «*двоїстий*» варіант властивості неперервність ймовірностей.

Властивість 9. (*Неперервність ймовірностей*). Припустимо, що маємо послідовність випадкових подій A_i, $i = 1, 2, \ldots, n, \ldots$, $(A_1, A_2, \ldots, A_n, \ldots \in \Im)$, для яких виконується наступна умова:

$$A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ \ldots$$

Тоді має місце наступна рівність:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \lim_{n\to\infty} P(A_n).$$

Властивість 10. Нехай $A_1, A_2, \ldots, A_n, \ldots \in \Im$ – довільні випадкові події. Завжди має місце наступна нерівність:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} P(A_i).$$

Доведення. Введемо послідовність випадкових подій $B_1, B_2, \ldots, B_n, \ldots \in \Im$ наступним чином:

$$B_1 = A_1,$$
$$B_2 = A_2 \cap (\Omega \backslash A_1),$$
$$B_3 = A_3 \cap (\Omega \backslash A_2) \cap (\Omega \backslash A_1),$$
$$\ldots$$
$$B_n = A_n \cap (\Omega \backslash A_{n-1}) \cap \ldots \cap (\Omega \backslash A_1),\ \ldots$$

Очевидно, що:

$$\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i \text{ , та } B_k \cap B_l = \varnothing \text{ , якщо } k \neq l.$$

Крім того, для будь-якого n: $B_n \subseteq A_n$. Тому на підставі аксіоми (**Р3**), та доведеної властивості 4, отримаємо:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Подібно, як і у випадку властивості 9, застосовуючи «*правила де-Моргана*», до властивість 10 отримаємо:

Властивість 11. Для довільних випадкових подій $A_1, A_2, ..., A_n, ... \in \Im$ завжди має місце наступна нерівність:

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) \geq 1 - \sum_{i=1}^{\infty} \left(1 - P(A_i)\right).$$