Конспект лекцій з теорії ймовірностей

Каніовська І.Ю.

19 липня 2020 р.

Зміст

Розділ 1

Випадкові події

1.1 Основні поняття теорії ймовірностей

1.1.1 Стохастичний експеримент (СЕ)

Означення 1.1.1. *Стохастичним експериментом (випробуванням)* називається експеримент, який можна повторювати неодноразово, зберігаючи певні умови, і результат якого експерименту заздалегідь передбачити неможливо.

Означення 1.1.2. Будь-який результат СЕ називається *подією*.

Приклад 1.1.1.

1. СЕ — кидання кубика один раз, подія — випало 6 очок.
2. СЕ — кидання кубика двічі, подія — сума очок, що випала, дорівнює 6.

Події бувають:

1. Випадкові — можуть відбутися чи не відбутися при проведенні СЕ.
2. Неможливі — ніколи не відбуваються в даному СЕ.
3. Вірогідні — завжди відбуваються в даному СЕ.

1.1.2 Теоретико-множинний підхід до основних понять ТЙ

Будемо вважати, що кожному СЕ можна поставити у відповідність деяку множину, що називається *простором елементарних подій* Ω. Під *елементарними подіями* ω будемо розуміти єдині логічно можливі події СЕ, що виключають одна одну.

Приклад 1.1.2.

1. СЕ — кидання кубика один раз. $\Omega = \{\omega_1, \omega_2, ..., \omega_6\}$, де $\omega_k = \{\text{випало } k \text{ очок}\}, k = 1, ..., 6$.
2. СЕ — кидання монетки до першої появи герба. $\Omega = \{\omega_1, \omega_2, ..., \omega_k, ...\}$,
   де $\omega_i = \{\text{герб випав на } i\text{-тому киданні}\}, i \in \mathbb{N}$.
3. СЕ — зустріч двох осіб, що домовилися зустрітися протягом години. x — час приходу першої особи, y — час приходу другої, $0 \le x, y \le 1$. $\Omega = \{(x, y) : 0 \le x, y \le 1\} \subset \mathbb{R}^2$.

В подальшому випадкові події позначатимемо $A, B, C, ...$. *Випадкова подія — підмножина* Ω. У прикладі з киданням кубика один раз $A = \{\text{випало 6 очок}\} = \{\omega_6\}$. В загальному випадку маємо $A = \{\omega_{k1}, \omega_{k2}, ..., \omega_{kn}, ...\} \subset \Omega$. Неможлива подія — $\varnothing$, вірогідна — Ω.

1.1.3 Основні операції над подіями

Зауваження. Розглядаємо події в межах фіксованих СЕ та простору елементарних подій.

1. *Включення* $A \subset B$ означає, що якщо відбулася подія A, то обов'язково відбудеться подія B. Наприклад, $A = \{\text{витягнуто даму пік}\}, B = \{\text{витягнуто карту чорної масті}\}$. Очевидно, $A \subset B$.

   Рівність подій: $(A \subset B, B \subset A) \iff (A = B)$.

2. *Об'єднання подій* $A \cup B$ — це подія, яка відбувається тоді, коли відбувається принаймні одна з подій A чи B.

   Властивості: $A \cup A = A$, $A \cup B = B \cup A$, $(A \subset B) \Rightarrow (A \cup B = B)$, $A \cup \Omega = \Omega$, $A \cup \varnothing = A$, $(A \cup B) \cup C = A \cup (B \cup C)$.

   Операція узагальнюється на скінченну або зліченну кількість подій: $A = \bigcup_{i=1}^{n(\infty)} A_i$.

3. *Перетин подій* $A \cap B$ — це подія, яка відбувається тоді, коли A і B відбуваються одночасно.

   Властивості: $A \cap A = A$, $A \cap B = B \cap A$, $(A \subset B) \Rightarrow (A \cap B = A)$, $A \cap \Omega = A$, $A \cap \varnothing = \varnothing$, $(A \cap B) \cap C = A \cap (B \cap C)$, $A \cap B \subset A$, $A \cap B \subset B$, $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.

   Операція узагальнюється на скінченну або зліченну кількість подій: $A = \bigcap_{i=1}^{n(\infty)} A_i$.

Означення 1.1.3. Події A та B називається *несумісними*, якщо вони не відбуваються одночасно: $A \cap B = \varnothing$. Узагальнення: події $A_1, A_2, ..., A_n, ...$ називаються *попарно несумісними*, якщо $A_i \cap A_j = \varnothing$ для $i \neq j$.

Означення 1.1.4. Події $A_1, A_2, ..., A_n, ...$ утворюють *повну групу подій*, якщо вони попарно несумісні та $\bigcup_{i=1}^{n(\infty)} A_i = \Omega$.

4. *Протилежна подія* $\overline{A}$ — це подія, яка відбувається тоді, коли A не відбувається.

   Властивості: $A \cup \overline{A} = \Omega$, $A \cap \overline{A} = \varnothing$, $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$, $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$.

5. *Різниця подій* $A \backslash B$ — це подія $A \cap \overline{B}$. Для протилежної події маємо $\overline{A} = \Omega \backslash A$.

1.1.4 Алгебра та σ-алгебра подій

Означення 1.1.5. Непорожня система підмножин $\mathcal{F}$ простору елементарних подій Ω утворює *алгебру подій*, якщо:

1. $\Omega \in \mathcal{F}$;
2. $(A, B \in \mathcal{F}) \Rightarrow (A \cup B \in \mathcal{F})$;
3. $(A \in \mathcal{F}) \Rightarrow \left(\overline{A} \in \mathcal{F}\right)$.

Узагальнення: якщо Ω містить зліченну кількість подій, то означення σ-алгебра отримаємо заміною умови 2 на $(\forall n \in \mathbb{N} : A_n \in \mathcal{F}) \Rightarrow (\bigcup_{n=1}^{\infty} A_n \in \mathcal{F})$.

Пара $\{\Omega, \mathcal{F}\}$ називається *вимірним простором стохастичного експерименту.*

1.1.5 Міра вірогідності появи випадкової події

Дослідника завжди цікавить кількісна характеристика появи тієї чи іншої події.

Нехай Ω — скінченний чи зліченний. Поставимо у відповідність кожній елементарній події ω_k число $p_k \geq 0$ так, що $\sum_{k=0}^{n(\infty)} p_k = 1$. Тоді $P(A) = \sum_{\omega_k \in A} p_k$ — кількісна характеристика, ймовірність події A.

Приклад 1.1.3. *Класична модель ймовірності.* Якщо простір елементарних подій Ω СЕ скінченний та всі ω_k рівноможливі, то такий СЕ називається *класичним.* В такому випадку $p_1 = p_2 = ... = p_n = \frac{1}{n}$, де $n = \text{card}(\Omega)$.

$$P(A) = \sum_{\omega_k \in A} p_k = \frac{m}{n} = \frac{\text{кількість елементарних подій в } A}{\text{загальна кількість елементарних подій}} \tag{1}$$

Ймовірності, що розраховуються за формулою (1), називаються *класичними*.

Приклад 1.1.4. 1. «Задачі про вибір» — коли з великої кількості чогось вибирають певну кількість. Наприклад, з урни з 10 кульками, 3 чорними та 7 білими, навмання витягають 5 кульок. Обчислимо ймовірність події $A = \{\text{серед них 2 чорних кульки}\}$: $P(A) = \frac{C_3^2 \cdot C_7^3}{C_{10}^5} = \frac{3 \cdot 35}{252} = \frac{5}{12}$.

2. «Задачі про ліфт». 5 осіб одночасно зайшли в ліфт 11-поверхового будинку. Яка ймовірність того, що вони всі вийдуть на різних поверхах, починаючи з другого?

   $P(A) = \frac{A_{10}^5}{\widetilde{A}_{10}^5} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{10^5} = 0.3024$. Тут A_n^k та $\widetilde{A}_n^k$ — кількості розміщень без повторень та з повтореннями відповідно.

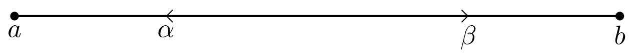
Властивості класичної ймовірності:

1. $\forall A \in \mathcal{F} : 0 \le P(A) \le 1$.
2. $P(\Omega) = 1$.
3. $P(\overline{A}) = 1 - P(A)$.
4. $P(\varnothing) = 0$.
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, для несумісних A і B $P(A \cup B) = P(A) + P(B)$.
6. $(A \subset B) \Rightarrow (P(A) \le P(B))$.
7. Якщо $A_1, A_2, ..., A_n$ — повна група подій СЕ, то $P(\bigcup_{i=1}^n A_i) = 1$.

1.2 Геометричні ймовірності. Аксіоми теорії ймовірностей.

1.2.1 Геометрична модель ймовірності

Приклад 1.2.1. Нехай точка кидається навмання на відрізок $[a; b]$. Яка ймовірність її потрапляння в $\langle \alpha; \beta \rangle \subset [a; b]$?
Розглянемо подію $A = \{\text{точка потрапила в } \langle \alpha; \beta \rangle\}$.

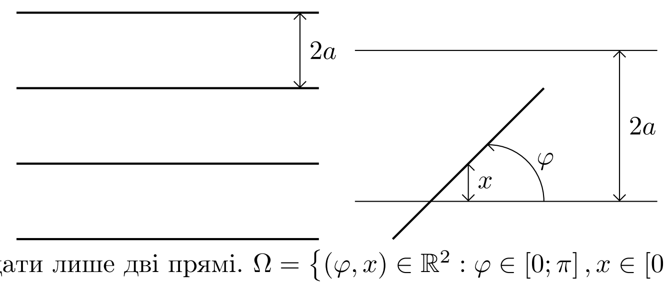


$P(A) = k \cdot l_{\langle \alpha; \beta \rangle}$ для деякого $k > 0$. З іншого боку, $P(\Omega) = 1 = k \cdot l_{[a;b]}$. Таким чином $k = \frac{1}{l_{[a;b]}} = \frac{1}{b-a}$. Тому $P(A) = \frac{l_{\langle \alpha; \beta \rangle}}{l_{[a;b]}}$.

Нехай простір елементарних подій інтерпретується як замкнена область в $\mathbb{R}^n$, а випадкові події — її підмножини. В якості σ-алгебри подій $\mathcal{F}$ беремо підмножини, що мають міру μ. Тоді в якості ймовірності деякої події A беремо $P(A) = \frac{\mu(A)}{\mu(\Omega)}$. Наприклад, в $\mathbb{R}^1$ беремо міру «довжина», в $\mathbb{R}^2$ — «площа», а в $\mathbb{R}^3$ — «об'єм».

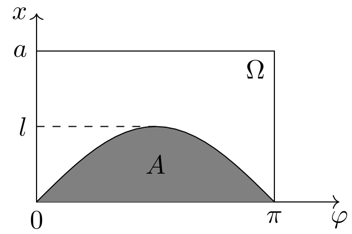
Ймовірності, що знаходяться таким чином називаються *геометричними*, а сама модель називається *геометричною моделлю ймовірності.* Геометрична модель може використовуватись, коли Ω має геометричну інтерпретацію як замкнена область в $\mathbb{R}^n$, а елементарні події — рівноможливі.

Приклад 1.2.2 (задача Бюффона)**.** На площині накреслені паралельні прямі на відстані $2a$, на них кидається голка довжиною $2l$, $l \le a$. Яка ймовірність того, що голка перетне якунебудь пряму?



Достатньо розглядати лише дві прямі. $\Omega = \left\{ (\varphi, x) \in \mathbb{R}^2 : \varphi \in [0; \pi] , x \in [0; a] \right\}$

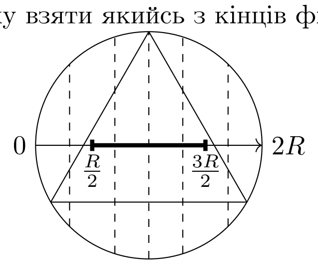
При такій побудові простору елементарних подій подія $A = \{\text{голка перетне пряму}\} = \{(\varphi, x) \in \Omega : x \leq l \sin\varphi\}$



$P(A) = \frac{S(A)}{S(\Omega)},\ S(\Omega) = \pi a,\ S(A) = \int_0^\pi l \sin\varphi \, \mathrm{d}x = l\,(-\cos\varphi)|_0^\pi = 2l$. Отже, $P(A) = \frac{2l}{\pi a}$.
Якщо провести n кидань голки, у m з яких голка потрапить на пряму, то за допомогою приблизної рівності $\frac{m}{n} \approx \frac{2l}{\pi a}$ можна приблизно обчислити число π. Наприклад, якщо взяти $n = 5000, m = 2532, l/a = 0.8$, то отримаємо $\pi \approx 3.1596$.

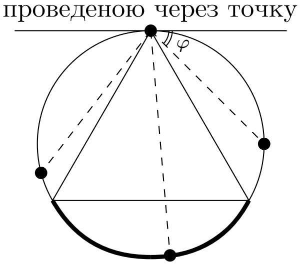
Приклад 1.2.3 (парадокс Бертрана)**.** Нехай в колі радіусом R навмання обирається хорда. Яка ймовірність того, що її довжина буде більшою за довжину сторони правильного трикутника, вписаного в це коло?
Існують три способи розв'язання цієї задачі, причому кожен з них дає різний результат.

Спосіб 1. З міркувань симетрії обирається якийсь фіксований діаметр кола і розглядається всі перпендикулярні до нього хорди. Серед них і обирається навмання довільна хорда. Очевидно, що кожна хорда у цьому випадку може бути однозначно визначена своєю серединою, тобто кожній хорді можна взаємно однозначно поставити у відповідність координату її середини, якщо за початок відліку взяти якийсь з кінців фіксованого діаметру.



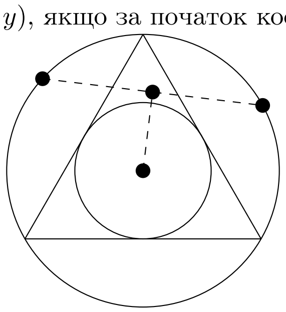
Таким чином, множина всіх значень координати середини хорди $\Omega = [0; 2R]$. Множина, що відповідає події — це відрізок $A = \left[\frac{R}{2}; \frac{3R}{2}\right]$. В якості міри візьмемо довжину. Тоді $P(A) = \frac{l(A)}{l(\Omega)} = \frac{R}{2R} = \frac{1}{2}$.

Спосіб 2. В цьому способі пропонується розглядати тільки хорди з одним закріпленим кінцем. Кожній хорді поставимо у відповідність ту частину дуги кола, яка потрапляє у кут φ, що утворює хорда з дотичною, проведеною через точку закріплення кінця всіх хорд.



Тоді множина $\Omega = [0; \pi]$, а $A = \left[\frac{\pi}{3}; \frac{2\pi}{3}\right]$. Знову візьмемо за міру довжину і отримаємо $P(A) = \frac{\pi/3}{\pi} = \frac{1}{3}$.

Спосіб 3. Розглядаються всі хорди кола. Кожній з них взаємно однозначно ставиться у відповідність точка її середини (x, y), якщо за початок координат прийняти центр кола.



В такому випадку $\Omega = \left\{(x,y) \in \mathbb{R}^2 : x^2+y^2 \le R^2\right\}$, а $A = \left\{(x,y) \in \Omega : x^2+y^2 \le \frac{R^2}{4}\right\}$. Тоді $P(A) = \frac{\pi R^2/4}{\pi R^2} = \frac{1}{4}$.

Жозеф Бертран був противником геометричного означення ймовірності. Він казав, що воно не дає можливості однозначно обчислити ймовірність однієї і тієї ж випадкової події та використовував вищенаведений приклад для доведення своєї правоти. Дійсно, було отримано три різних відповіді при розв'язанні однієї задачі. Так в чому ж справа? Насправді, помилка полягає у різних трактуваннях поняття «навмання обрана хорда». В кожному з трьох способів це трактування було різним, що й стало причиною різних відповідей.

1.2.2 Аксіоми теорії ймовірностей

I. Побудова вимірного простору $\{\Omega, \mathcal{F}\}$:

A1: Будь-якому стохастичному експерименту можна поставити у відповідність простір елементарних подій Ω.

A2: $(\forall n \in \mathbb{N} : A_n \in \mathcal{F}) \Rightarrow \left(\bigcup_{n=1}^{n(\infty)} A_n \in \mathcal{F}\right)$.

A3: $(A \in \mathcal{F}) \Rightarrow \left(\overline{A} \in \mathcal{F}\right)$.
Дві останні аксіоми стосуються побудови алгебри чи σ-алгебри подій.

II. Властивості ймовірності:

P1: $\forall A \in \mathcal{F} : P(A) \ge 0$.

P2: $P(\Omega) = 1$ — аксіома нормування.

P3: $\forall A_1, A_2, ..., A_n, ... \in \mathcal{F}, A_i \cap A_j$ при $i \ne j$: $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ — аксіома зліченної адитивності. Іноді у випадках «простої» алгебри (а не σ-алгебри) достатньо аксіоми скінченної адитивності **P3′**: $\forall A_1, A_2, ..., A_n \in \mathcal{F}, A_i \cap A_j$ при $i \ne j$: $P(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k)$.

Означення 1.2.1. Нормована міра P, що введена на вимірному просторі $\{\Omega, \mathcal{F}\}$, називається *ймовірністю*. Всі події беруться з σ-алгебри подій. Трійка $\{\Omega, \mathcal{F}, P\}$ називається *ймовірнісний простором* стохастичного експерименту.

Ця система аксіом несуперечна, бо існують стохастичні експерименти, які задовольняють цим аксіомам, але неповна, бо в різних задачах теорії ймовірностей розглядаються різні ймовірнісні простори.

1.2.3 Властивості ймовірності, що випливають з аксіом

1. Якщо $A_1, A_2, ..., A_n, ... \in \mathcal{F}$ утворюють повну групу подій, то $P(\cup_{k=1}^{\infty} A_k) = 1$.

   Доведення. Випливає з аксіоми **P2**. ▲

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

   Доведення. $A \cup \overline{A} = \Omega \Rightarrow 1 \overset{P2}{=} P(\Omega) = P(A \cup \overline{A}) \overset{P3'}{=} P(A) + P(\overline{A})$. ▲

3. $P(\varnothing) = 0$.

   Доведення. Наслідок з властивості 2 ($\varnothing = \overline{\Omega}$). ▲

   Зауваження. Якщо ймовірність події дорівнює нулю, то це не означає, що подія неможлива.

   Приклад 1.2.4. СЕ — кидання точки на відрізок $[a; b]$. А = {точка потрапила в певну $x \in [a; b]$}. A не є неможливою, проте $P(A) = 0$.

4. $A \subset B \Rightarrow P(A) \le P(B)$.

Доведення. $A \subset B \Rightarrow B = A \cup (B \backslash A) \overset{P3'}{\Rightarrow} P(B) = P(A) + P(B \backslash A) \overset{P1}{\geq} P(A)$. ▲

5. $\forall A \in \mathcal{F} : P(A) \leq 1$.

   Доведення. Наслідок з властивості 4, $A \subset \Omega$ та **P2**. ▲

6. $A \subset B : P(B \backslash A) = P(B) - P(A)$.

   Доведення. Наслідок з доведення властивості 4. ▲

7. $\forall A, B \in \mathcal{F} : P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

   Доведення. $A \cup B = (A \backslash (A \cap B)) \cup (A \cap B) \cup (B \backslash (A \cap B))$ — попарно несумісні події. З аксіоми **P3′**: $P(A \cup B) = P(A \backslash (A \cap B)) + P(A \cap B) + P(B \backslash (A \cap B)) \overset{6}{=} P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) = P(A) + P(B) - P(A \cap B)$. ▲

8. Узагальнення властивості 7 — формула включення-виключення для ймовірностей: $\forall A_1, A_2, \ldots, A_n \in \mathcal{F} : P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i<j}^n P(A_i \cap A_j) + \sum_{i<j<k}^n P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} P(\bigcap_{i=1}^n A_i)$.

   Вправа. Довести.

Приклад 1.2.5 (задача про неуважну секретарку)**.** Секретарка поклала n листів в n чистих конвертів, заклеїла ці конверти і тільки після цього написала адреси. Яка ймовірність того, що хоча б один з листів дійде за призначенням?
$A_i = \{i\text{-тий лист дійшов за призначенням}\}, i = 1, ..., n$. $P(A_i) = \frac{1}{n} = \frac{(n-1)!}{n!}$.
$A = \{\text{хоча б один із листів дійшов за призначенням}\}, A = \bigcup_{i=1}^n A_i$.
$P(A_i \cap A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, i \neq j$.
$P(A_i \cap A_j \cap A_k) = \frac{1}{n(n-1)(n-2)}, i \neq j \neq k$.
...
$P(A_1 \cap \ldots \cap A_n) = \frac{1}{n!}$.
За формулою включення-виключення $P(A) = n \cdot \frac{1}{n} - C_n^2 \cdot \frac{1}{n(n-1)} - \cdots + (-1)^{n-1} \frac{1}{n!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1} \cdot \frac{1}{n!} \approx 1 - \frac{1}{e} \approx 0.63$.

9. $\forall A_1, A_2, ..., A_n, ... \in \mathcal{F} : P(\bigcup_{k=1}^\infty A_k) \leq \sum_{k=1}^\infty P(A_k)$

   Доведення. Введемо події $B_1 = A_1, B_2 = \overline{A_1} \cap A_2, B_3 = \overline{A_1} \cap \overline{A_2} \cap A_3, ..., B_n = \overline{A_1} \cap \overline{A_2} \cap \ldots \cap \overline{A_{n-1}} \cap A_n$. Ці події попарно несумісні, $\bigcup_{k=1}^\infty B_k = \bigcup_{k=1}^\infty A_k,\ B_k \subset A_k$ для всіх $k \in \mathbb{N}$.
   $P(\bigcup_{k=1}^\infty A_k) = P(\bigcup_{k=1}^\infty B_k) \overset{P3}{=} \sum_{k=1}^\infty P(B_k) \overset{4}{\leq} \sum_{k=1}^\infty P(A_k)$. ▲

10. $\forall A_1, A_2, ..., A_n, ... \in \mathcal{F} : P(\bigcap_{k=1}^\infty A_k) = 1 - P(\overline{\bigcap_{k=1}^\infty A_k}) = 1 - P(\bigcup_{k=1}^\infty \overline{A_k}) \geq 1 - \sum_{k=1}^\infty P(\overline{A_k})$.

1.2.4 Теореми неперервності ймовірності

Теорема 1. *Нехай є монотонно неспадна послідовність подій $A_1 \subset A_2 \subset \ldots \subset A_n \ldots \in \mathcal{F}$. Тоді $P(\bigcup_{n=1}^\infty A_n) = \lim_{n\to\infty} P(A_n)$.*

Доведення. $\bigcup_{n=1}^\infty A_n = A_1 \cup (A_2 \backslash A_1) \cup (A_3 \backslash A_2) \cup \ldots \cup (A_n \backslash A_{n-1}) \cup \ldots$
З аксіоми **P3** $P(\bigcup_{n=1}^\infty A_n) = P(A_1) + \sum_{n=2}^\infty P(A_n \backslash A_{n-1})$.
$S_n = P(A_1) + \sum_{k=2}^n P(A_k \backslash A_{k-1}) = P(A_1) + P(A_2) - P(A_1) + P(A_3) - P(A_2) + \ldots + P(A_n) - P(A_{n-1}) = P(A_n)$. Отже, $P(\bigcup_{n=1}^\infty A_n) = \lim_{n\to\infty} S_n = \lim_{n\to\infty} P(A_n)$. ▲

Теорема 2. *Нехай є монотонно спадна послідовність подій $A_1 \supset A_2 \supset \ldots \supset A_n \ldots \in \mathcal{F}$. Тоді $P(\bigcap_{n=1}^\infty A_n) = \lim_{n\to\infty} P(A_n)$.*

Доведення. $P(\bigcap_{n=1}^\infty A_n) = 1 - P(\bigcup_{n=1}^\infty \overline{A_n}) \overset{\text{теор. 1}}{=} 1 - \lim_{n\to\infty} P(\overline{A_n}) = \lim_{n\to\infty}(1 - P(\overline{A_n})) = \lim_{n\to\infty} P(A_n)$. ▲

1.3 Аксіома неперервності. Умовні ймовірності. Незалежність подій. Формула Баєса.

1.3.1 Аксіома неперервності

Теорема 3. *Аксіома зліченної адитивності **P3** еквівалентна аксіомі неперервності **P4**:*

$$\forall n \in \mathbb{N} : B_n \in \mathcal{F}, B_1 \supset B_2 \supset ... \supset B_n \supset ..., \bigcap_{n=1}^{\infty} B_n = \varnothing \Rightarrow \lim_{n\to\infty} P(B_n) = 0$$

Доведення. За теоремою 2 маємо наслідок **P3** $\Rightarrow$ **P4**.

Доведемо **P4** $\Rightarrow$ **P3**. Для подій $A_1, A_2, ..., A_n, ... \in \mathcal{F}, A_i \cap A_j$ при $i \neq j$ введемо події $B_n = \bigcup_{k=n+1}^{\infty} A_k$. Для них виконується $B_1 \supset B_2 \supset ... \supset B_n \supset ...$ та $\bigcap_{n=1}^{\infty} B_n = \varnothing$, тому з аксіоми **P4** $\lim_{n\to\infty} P(B_n) = 0$.

З іншого боку, маємо таку рівність:

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = P\left(\bigcup_{k=1}^{n} A_k\right) + P\left(\bigcup_{k=n+1}^{\infty} A_k\right) = \sum_{k=1}^{n} P(A_k) + P(B_n)$$

Перейшовши у правій частині до границі при $n \to \infty$, отримаємо $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$, що є твердженням аксіоми зліченної адитивності. ▲

1.3.2 Умовні ймовірності

Припустимо, що спостерігається деякий експеримент, що описується класичною моделлю, а для події B $P(B) > 0$. Умовна ймовірність $P(A/B)$ — це ймовірність події A за умови, що відбулась подія B. Наприклад, експеримент — витягання карт з колоди, $A = \{$витягнуто даму пік$\}$, $B = \{$витягнуто карту чорної масті$\}$.

Оскільки експеримент описується класичною моделлю, то можемо позначити m_B та $m_{A\cap B}$ кількості елементарних подій, що сприяють появами подій B та $A \cap B$ відповідно. Тоді $P(A/B) = \frac{m_{A\cap B}}{m_B} = \frac{m_{A\cap B}/n}{m_B/n} = \frac{P(A\cap B)}{P(B)}$, де $n = \text{card}(\Omega)$.

Ця формула справджується й для довільних ймовірнісних просторів.

Означення 1.3.1. Якщо подія B має додатну ймовірність $P(B) > 0$, то *умовна ймовірність* події A за умови, що відбулась подія B, обчислюється за формулою

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \tag{1}$$

Властивості умовної ймовірності:

1. $P(A/B) \geq 0$.
2. $P(\Omega/B) = P(B/B) = 1$.
3. $\forall A_1, A_2, ..., A_n, ... \in \mathcal{F} : A_i \cap A_j$ при $i \neq j :$
   $P((\bigcup_{n=1}^{\infty} A_n)/B) = \frac{P((\bigcup_{n=1}^{\infty} A_n)\cap B)}{P(B)} = \frac{P(\bigcup_{n=1}^{\infty}(A_n\cap B))}{P(B)} \overset{P3}{=} \frac{\sum_{n=1}^{\infty} P(A_n\cap B)}{P(B)} = \sum_{n=1}^{\infty} P(A_n/B)$.

Таким чином, для умовної ймовірності діють всі аксіоми ТЙ. Фактично, розглядається звуження ймовірнісного простору $\{B, \mathcal{F}_B, P(\cdot/B)\}$, в якому B стає вірогідною подією.

Теореми множення

З формули (1) випливає, що

$$P(A \cap B) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B) \tag{2}$$

Для трьох подій: $P(A \cap B \cap C) = P(A \cap B) \cdot P(C/(A \cap B)) = P(A) \cdot P(B/A) \cdot P(C/(A \cap B))$. За методом математичної індукції неважко довести, що

$$P\left(\bigcap_{n=1}^{n} A_n\right) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/(A_1 \cap A_2)) \cdot ... \cdot P\left(A_n/\left(\bigcap_{n=1}^{n-1} A_n\right)\right) \tag{3}$$

Приклад 1.3.1. На 10 картках написано по одній букві так, що можна скласти слово «математика». Дитина навмання витягує без повернення по одній картці. Яка ймовірність того, що після витягання 4 карток утвориться слово «мама»?

Треба обчислити ймовірність $P(M^{(1)} \cap A^{(2)} \cap M^{(3)} \cap A^{(4)})$, де верхній індекс означає крок, на якому витягнуто літеру. $P(M^{(1)}) = \frac{2}{10}$, $P(A^{(2)}/M^{(1)})\frac{3}{9}$, $P(M^{(3)}/(M^{(1)} \cap A^{(2)})) = \frac{1}{8}$, $P(A^{(4)}/(M^{(1)} \cap A^{(2)} \cap M^{(3)})) = \frac{2}{7}$. Тому за формулою (3) $P(M^{(1)} \cap A^{(2)} \cap M^{(3)} \cap A^{(4)}) = \frac{2 \cdot 3 \cdot 1 \cdot 2}{10 \cdot 9 \cdot 8 \cdot 7}$.

1.3.3 Незалежність подій

Означення 1.3.2. Дві події A та B називаються *незалежними*, якщо

$$P(A \cap B) = P(A) \cdot P(B) \tag{4}$$

Властивості незалежних подій:

1. Якщо події A та B незалежні, то $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A)$.

   Аналогічно $P(B/A) = P(B)$.

   Зауваження. Несумісні події залежні, якщо їх ймовірності не нульові.

2. Якщо події A та B незалежні, то пари A та $\overline{B}$, $\overline{A}$ та B, $\overline{A}$ та $\overline{B}$ — теж незалежні.

   Доведення. Доведемо спочатку для пари A та $\overline{B}$. $P(A \cap \overline{B}) = P(A \backslash (A \cap B)) = P(A) - P(A \cap B) = P(A) - P(A) \cdot P(B) = P(A) \cdot (1 - P(B)) = P(A) \cdot P(\overline{B})$. Отже, A та $\overline{B}$ — незалежні. Доведення для інших пар аналогічне. ▲

3. Якщо A та B, A та C незалежні, а B та C несумісні, то A та $B \cup C$ — незалежні події.

   Доведення. $P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C)) = P(A \cap B) + P(A \cap C) = P(A) \cdot P(B) + P(A) \cdot P(C) = P(A) \cdot (P(B) + P(C)) = P(A) \cdot P(B \cup C)$. ▲

1.3.4 Незалежність у сукупності

Означення 1.3.3. Події $A_1, A_2, ..., A_n \in \mathcal{F}$ називаються *незалежними у сукупності*, якщо

$$\forall i_1, i_2, ... i_r \in 1, ..., n : P\left(\bigcap_{k=1}^{r} A_{i_k}\right) = \prod_{k=1}^{r} P\left(A_{i_k}\right) \tag{5}$$

Зокрема:

1. $\forall i, j \in 1, .., n, i \neq j : P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$ — попарна незалежність.
2. $P(\bigcap_{k=1}^{n} A_k) = \prod_{k=1}^{n} P(A_k)$.

Зауваження. З незалежності подій у сукупності випливає попарна незалежність, але наслідку в іншу сторону, взагалі кажучи, немає.

Приклад 1.3.2 (приклад Бернштейна)**.** Дитина кидає на підлогу тетраедр, три грані якого розфарбовані відповідно зеленим, синім та червоним кольором, а четверта — усіма трьома кольорами. Введемо події З, С, Ч, які означають, що випала грань, на якій є зелений, синій або червоний колір відповідно.

$P(\text{З}) = P(\text{С}) = P(\text{Ч}) = \frac{1}{2}$, $P(\text{З} \cap \text{С}) = P(\text{З} \cap \text{Ч}) = P(\text{С} \cap \text{Ч}) = \frac{1}{4}$, тому ці події є попарно незалежними. Але $P(\text{З} \cap \text{С} \cap \text{Ч}) = \frac{1}{4} \neq \frac{1}{8}$, тому незалежності у сукупності немає.

Теорема додавання для незалежних у сукупності подій. $A_1, A_2, ..., A_n \in \mathcal{F}$ — незалежні у сукупності. $P(\bigcup_{k=1}^{n} A_k) = ?$

$P(\bigcup_{k=1}^{n} A_k) = 1 - P(\overline{\bigcup_{k=1}^{n} A_k}) = 1 - P(\bigcap_{k=1}^{n} \overline{A_k}) = 1 - \prod_{k=1}^{n} P(\overline{A_k})$.
Зокрема, для двох незалежних подій A і B: $P(A \bigcup B) = 1 - P(\overline{A}) \cdot P(\overline{B})$

1.3.5 Формула повної ймовірності

Нехай деякому СЕ поставлено у відповідність простір елементарних подій Ω.

$H_1, H_2, ..., H_n \in \mathcal{F}$ (може бути й зліченна кількість) — повна група подій, які називаємо *гіпотезами (припущеннями)*. Нехай подія A відбувається разом з якоюсь гіпотезою. Тоді має місце *формула повної ймовірності*:

$$P\left(A\right) = \sum_{k=1}^{n} P(H_k) \cdot P(A/H_k) \tag{6}$$

Доведення. $P(A) = P(A \cap \Omega) = P(A \cap (\bigcup_{k=1}^{n} H_k)) = P((\bigcup_{k=1}^{n}(A \cap H_k))) = \sum_{k=1}^{n} P(A \cap H_k) = \sum_{k=1}^{n} P(H_k) \cdot P(A/H_k)$. ▲

Приклад 1.3.3.

1. У магазин постачають 80% телефонів з Китаю, 15% з В'єтнаму та 5% з Кореї, причому бракованих відповідно 1%, 0.1% та 0.01%. Знайти ймовірність події $A = \{$куплений телефон буде бракованим$\}$.

   За умовою введемо гіпотези $H_1 = \{$телефон з Китаю$\}$, $H_2 = \{$телефон з В'єтнаму$\}$ та $H_3 = \{$телефон з Кореї$\}$. $P(H_1) = 0.8$, $P(H_2) = 0.15$, $P(H_3) = 0.05$, $P(A/H_1) = 0.01$, $P(A/H_2) = 0.001$, $P(A/H_3) = 0.0001$. Тоді за формулою (6) $P(A) = 0.8 \cdot 0.01 + 0.15 \cdot 0.001 + 0.05 \cdot 0.0001 = 0.080155$.

2. Серед N екзаменаційних білетів n «щасливих», $n < N$. У якого студента ймовірність витягнути «щасливий» білет більша — у того, хто тягне першим, чи того, хто тягне другим?

   Позначимо через A і B події, що «щасливий» білет витягнув перший та другий студент відповідно. За умовою $P(A) = \frac{n}{N}$. Щоб скористатися формулою (6) для обчислення $P(B)$, введемо гіпотези $H_1 = A$ та $H_2 = \overline{A}$. $P(H_1) = \frac{n}{N}$, $P(H_2) = \frac{N-n}{N}$.
   $P(B) = P(H_1) \cdot P(B/H_1) + P(H_2) \cdot P(B/H_2) = \frac{n}{N} \cdot \frac{n-1}{N-1} + \frac{N-n}{N} \cdot \frac{n}{N-1} = \frac{n^2-n+N \cdot n-n^2}{N \cdot (N-1)} = \frac{n \cdot (N-1)}{N \cdot (N-1)} = \frac{n}{N}$.

1.3.6 Формула Баєса

Як і раніше, нехай $H_1, H_2, ..., H_n \in \mathcal{F}$ — повна група подій деякого СЕ, які називаємо гіпотезами, причому перед проведенням експерименту відомі їх *апріорні* ймовірності $P(H_1), P(H_2), ..., P(H_n)$. В результаті проведення експерименту відбулась деяка подія A. Постає питання: чому рівні *апостеріорні* ймовірності $P(H_1/A), P(H_2/A), ..., P(H_n/A)$? Тобто, чому дорівнюють ймовірності, що мала місце кожна з гіпотез за умови, що подія A відбулась? На це дає відповідь *формула Баєса*:

$$P(H_i/A) = \frac{P(H_i \cap A)}{P(A)} = \frac{P(H_i) \cdot P(A/H_i)}{\sum_{k=1}^{n} P(H_k) \cdot P(A/H_k)}, i = 1, ..., n \tag{7}$$

Приклад 1.3.4. В групі з 10 осіб троє вчаться на «5», четверо на «4», двоє на «3» та один на «2». Викладач підготував на екзамен 20 питань. Студенти, які вчаться на «5», знають відповіді на всі, на «4» — на 16, на «3» — на 10, на «2» — на 5. Екзаменаційний білет містить 3 питання. Деякий студент відповів правильно на всі 3. Яка ймовірність того, що він вчиться на «2»?

Введемо гіпотези $H_1 = \{$студент вчиться на «5»$\}$, $H_2 = \{$студент вчиться на «4»$\}$, $H_3 = \{$студент вчиться на «3»$\}$, $H_4 = \{$студент вчиться на «2»$\}$. За умовою $P(H_1) = \frac{3}{10}$, $P(H_2) = \frac{4}{10}$, $P(H_3) = \frac{2}{10}$, $P(H_4) = \frac{1}{10}$. Позначимо $A = \{$студент відповів на всі три питання$\}$. Тоді $P(A/H_1) = 1$, $P(A/H_2) = \frac{16 \cdot 15 \cdot 14}{20 \cdot 19 \cdot 18}$, $P(A/H_3) = \frac{10 \cdot 9 \cdot 8}{20 \cdot 19 \cdot 18}$, $P(A/H_4) = \frac{5 \cdot 4 \cdot 3}{20 \cdot 19 \cdot 18}$. За формулою (7) шукана ймовірність $P(H_4/A) = \frac{P(H_4) \cdot P(A/H_4)}{\sum_{k=1}^{4} P(H_k) \cdot P(A/H_k)} = \frac{0.1 \cdot 5 \cdot 4 \cdot 3}{0.3 \cdot 20 \cdot 19 \cdot 18 + 0.4 \cdot 16 \cdot 15 \cdot 14 + 0.2 \cdot 10 \cdot 9 \cdot 8 + 0.1 \cdot 5 \cdot 4 \cdot 3} = \frac{60}{35460} = \frac{1}{591}$.

1.4 Поліноміальна модель ймовірності. Схема Бернуллі.

1.4.1 Поліноміальна схема

Припустимо, що події $A_1, A_2, ..., A_k$ утворюють повну групу подій деякого СЕ, причому відомі ймовірності $P(A_i) = p_i, \sum_{i=1}^{k} p_i = 1$. Експеримент проводиться n разів, в кожному з яких може відбутись одна з подій A_i. Можливими результатами n раз проведеного експерименту будуть події $A_{i_1}^{(1)} \cap A_{i_2}^{(2)} \cap ... \cap A_{i_n}^{(n)}$, де $A_{i_r}^{(r)}, r = 1, ..., n, i_r = 1, ..., k$ — один із можливих результатів r-того випробування.

Означення 1.4.1. Випробування, що проводяться, називаються *незалежними*, якщо

$$P\left(A_{i_1}^{(1)} \cap A_{i_2}^{(2)} \cap ... \cap A_{i_n}^{(n)}\right) = P\left(A_{i_1}^{(1)}\right) \cdot P\left(A_{i_2}^{(2)}\right) \cdot ... \cdot P\left(A_{i_n}^{(n)}\right)$$

Основна задача — знайти ймовірність того, що подія A_1 відбудеться m_1 разів, подія A_2 відбудеться m_2 разів і т.д., подія A_k відбудеться m_k разів, причому $\sum_{j=1}^{k} m_j = n$. Такі ймовірності будемо позначати як $P_n(m_1, m_2, ..., m_n)$.

$$P\left(\underbrace{A_1 \cap ... \cap A_1}_{m_1 \text{разів}} \cap \underbrace{A_2 \cap ... \cap A_2}_{m_2 \text{разів}} \cap ... \cap \underbrace{A_k \cap ... \cap A_k}_{m_k \text{разів}}\right) = p_1^{m_1} p_2^{m_2} ... p_k^{m_k}$$

Це є лише один із можливих варіантів послідовності появи подій в серії випробувань. Всього таких варіантів рівно $C_n^{m_1, m_2, ..., m_k} = \frac{n!}{m_1! m_2! ... m_k!}$. Таким чином отримуємо:

$$P_n(m_1, ..., m_k) = \frac{n!}{m_1! m_2! ... m_k!} p_1^{m_1} p_2^{m_2} ... p_k^{m_k} \tag{1}$$

Означення 1.4.2. Ймовірності, що обчислюємо за даною формулою, називаються *поліноміальними*, а сама схема — *поліноміальною схемою ймовірності.*

Зауваження. Звідки назва «поліноміальна»?

$$1 = (p_1 + p_2 + ... + p_k)^n = \sum_{m_1 + ... + m_k = n} P_n(m_1, ..., m_k)$$

Приклад 1.4.1. Студент ІПСА за рівнем підготовки з ймовірністю 0.3 вважається слабким студентом, з ймовірністю 0.5 вважається середнім студентом та з ймовірністю 0.2 — сильним студентом. Яка ймовірність того, що з 6 навмання обраних студентів кількість слабких та сильних буде однаковою?

Розглядаємо подію $A = \{\text{кількість слабких рівна кількості сильних}\}$. Позначимо ймовірності того, що студент має певний рівень підготовки, таким чином: $p_1 = 0.3$, $p_2 = 0.5$, $p_3 = 0.2$. $P(A) = ?$

Скористаємось поліноміальною схемою: $P(A) = P_6(3, 0, 3) + P_6(2, 2, 2) + P_6(1, 4, 1) + P_6(0, 6, 0) = \frac{6!}{3!3!} \cdot (0.3)^3 \cdot (0.2)^3 + \frac{6!}{2!2!2!} \cdot (0.3)^2 \cdot (0.5)^2 \cdot (0.2)^2 + \frac{6!}{1!4!1!} \cdot 0.3 \cdot (0.5)^4 \cdot 0.2 + \frac{6!}{6!} \cdot (0.5)^6 = 0.213445.$

1.4.2 Схема Бернуллі

Означення 1.4.3. Поліноміальна схема, в кожному випробуванні якої є тільки дві події A_1 та A_2, називається *біноміальною схемою* або *схемою Бернуллі.*

A_1 та A_2 утворюють повну групу подій $\Rightarrow A_2 = \overline{A_1}$. Позначимо $A_1 = A$ — «успіх», $A_2 = \overline{A}$ — «невдача», $P(A) = p, P(\overline{A}) = 1 - p = q$.

Яка ймовірність того, що в n випробуваннях «успіх» з'явиться m разів? На це питання дає відповідь *формула Бернуллі*:

$$P_n(m) = \frac{n!}{m!(n-m)!} p^m q^{n-m} = C_n^m p^m q^{n-m} \tag{2}$$

Ймовірності, що обчислюються за цією формулою, називаються *біноміальними.*

Приклад 1.4.2. Спортсмен 5 разів стріляє по мішені. Ймовірність влучення при кожному пострілі — 0.2. Яка ймовірність того, що він влучив не менше 3 разів?

$P\{\text{влучив} \geq 3 \text{ разів}\} = P_5(3) + P_5(4) + P_5(5) = C_5^3 (0.2)^3 (0.8)^2 + C_5^4 (0.2)^4 0.8 + C_5^5 (0.2)^5 = 0.05792.$

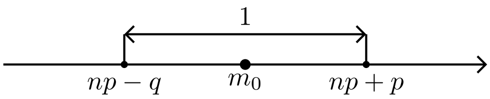
1.4.3 Найбільш імовірна кількість успіхів в схемі Бернуллі

Означення 1.4.4. Натуральне число m_0, при якому $P_n(m)$ набуває найбільшого значення, називається *найбільш імовірною кількістю успіхів.*

Розглянемо систему:

$$\begin{cases} P_n(m_0-1) \le P_n(m_0) \\ P_n(m_0+1) \le P_n(m_0) \end{cases} \Leftrightarrow \begin{cases} \frac{n!}{(m_0-1)!(n-m_0+1)!} p^{m_0-1} q^{n-m_0+1} \le \frac{n!}{m_0!(n-m_0)!} p^{m_0} q^{n-m_0} \\ \frac{n!}{(m_0+1)!(n-m_0-1)!} p^{m_0+1} q^{n-m_0-1} \le \frac{n!}{m_0!(n-m_0)!} p^{m_0} q^{n-m_0} \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \frac{q}{n-m_0+1} \le \frac{p}{m_0} \\ \frac{p}{m_0+1} \le \frac{q}{n-m_0} \end{cases} \Leftrightarrow \begin{cases} m_0 \le np+p \\ m_0 \ge np-q \end{cases}$$



Таким чином отримуємо два варіанти:

1. $np+p$ – не ціле $\Rightarrow m_0 = [np+p]$
2. $np+p$ – ціле $\Rightarrow \begin{cases} m_0^{(1)} = np+p \\ m_0^{(2)} = np-q \end{cases}$

Приклад 1.4.3. Знайти найбільш імовірну кількість успіхів для прикладу 1.4.2.

$np+p = 1.2 \Rightarrow m_0 = [np+p] = [1.2] = 1$. Ймовірність лише одного влучення буде найбільшою, рівною 0.904.

1.4.4 Асимптотичні наближення формули Бернуллі

Формула Бернуллі не є зручною для великих значень m та n. Існують формули, якими можна скористатись для наближення результатів формули Бернуллі.

Твердження 4 (асимптотична формула Пуассона)**.** *Нехай n — достатньо велике натуральне число, а p — достатньо мале, так що $a = np \in (1, 20)$. Тоді для наближення формули Бернуллі можна використати асимптотичну формулу Пуассона:*

$$P_n(m) \approx \frac{(np)^m}{m!} e^{-np}$$

Доведення (нестроге). $P_n(m) = \frac{n(n-1)\ldots(n-m+1)}{m!} \left(\frac{a}{n}\right)^m (1-\frac{a}{n})^{n-m} = \frac{a^m}{m!} 1 (1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{m+1}{n})(1-\frac{a}{n})^n (1-\frac{a}{n})^{-m} \underset{n\to\infty}{\longrightarrow} \frac{a^m}{m!} e^{-a} = \frac{(np)^m}{m!} e^{-np}$. ▲

Твердження 5 (формула Муавра-Лапласа)**.** *Нехай $npq \ge 20$. Тоді для наближення формули Бернуллі можна використати формулу Муавра-Лапласа:*

$$P_n(m) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(m-np)^2}{2npq}}$$

Доведення. Грунтується на локальній теоремі Муавра-Лапласа. ▲