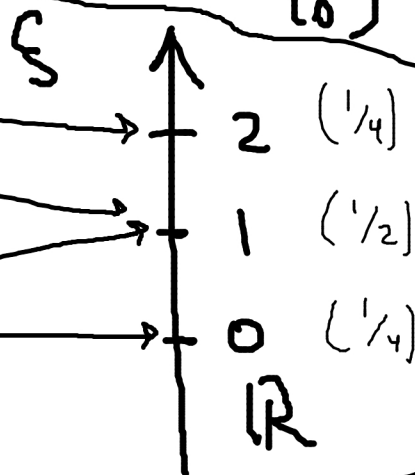
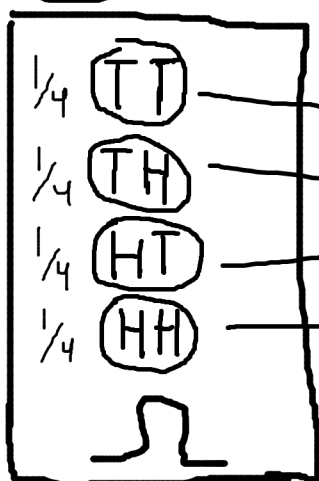


$(\Omega, \mathcal{F}, \mathbb{P})$

Probability Theory

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(B/A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

$\mathbb{E}\xi$   
 $\mathbb{D}\xi$   
 $G_\xi$

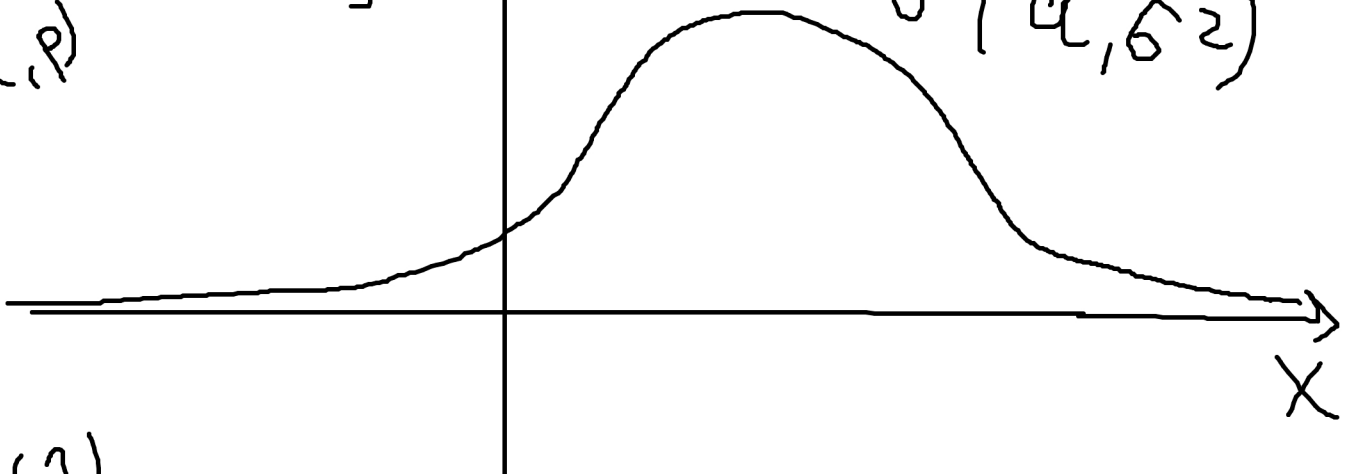


Throwing a coin twice  
 $\xi$  - the amount of tails (T).

$\text{Bin}(n, p)$

$f_\xi(x)$

$\xi \sim \mathcal{N}(\mu, \sigma^2)$



$\text{Pois}(\lambda)$   
 $\text{Exp}(\lambda)$

# Зміст

<b>1 Основи теорії ймовірності</b>	<b>4</b>
1.1 Імовірнісний простір	4
1.2 Класична модель ймовірності	5
1.3 Геометрична ймовірність	6
1.4 Умовна ймовірність	8
1.5 Незалежність подій	9
1.6 Формула повної ймовірності	10
1.7 Формула Баєса	11
1.8 Схема Бернуллі	11
1.9 Найімовірніша кількість успіхів у схемі Бернуллі	12
1.10 Поліноміальна схема	13
<b>2 Випадкові величини</b>	<b>14</b>
2.1 Перші означення	14
2.2 Основні види випадкових величин	16
2.2.1 Дискретна випадкова величина	16
2.2.2 Абсолютно неперервна випадкова величина	16
2.3 Числові характеристики випадкових величин	19
2.3.1 Математичне сподівання	19
2.3.2 Дисперсія, середньоквадратичне відхилення	20
2.4 Коротко про генератриси в теорії ймовірностей	21
2.5 Популярні розподіли дискретних випадкових величин	22
2.5.1 Розподіл Бернуллі	22
2.5.2 Біноміальний розподіл	22
2.5.3 Геометричний розподіл	23
2.5.4 Розподіл Пуассона	23
2.6 Популярні розподіли абсолютно неперервних випадкових величин	26
2.6.1 Рівномірний розподіл	26
2.6.2 Експоненційний розподіл	26
2.6.3 Нормальний розподіл	27
<b>3 Випадкові вектори</b>	<b>29</b>
3.1 Перші означення	29
3.2 Основні види випадкових векторів	29
3.2.1 Дискретний випадковий вектор	29
3.2.2 Абсолютно неперервний випадковий вектор	30
3.3 Числові характеристики випадкових векторів	31
3.3.1 Математичне сподівання	31
3.3.2 Коваріація (кореляційний момент)	31
3.4 Коваріаційна матриця	32
3.5 Рівномірний розподіл на площині	32
<b>4 Функції від випадкових величин та векторів</b>	<b>33</b>
4.1 Одновимірна функція	33
4.1.1 Дискретна випадкова величина	33
4.1.2 Абсолютно неперервна випадкова величина	33
4.2 Багатовимірна функція	34
4.2.1 Дискретний випадковий вектор	34
4.2.2 Абсолютно неперервний випадковий вектор	35
4.3 Щільності розподілу суми, добутку та частки випадкової величини	36
4.3.1 $\xi_1 + \xi_2$	36
4.3.2 $\xi_1 \cdot \xi_2$	36
4.3.3 $\frac{\xi_2}{\xi_1}$	37
4.4 Розподіл максимуму та мінімуму	38
4.5 Математичне сподівання функцій від випадкових величин/векторів	38

<b>5</b>	<b>Характеристичні функції</b>	<b>40</b>
5.1	Характеристичні функції випадкових величин . . . . .	40
<b>6</b>	<b>Граничні теореми</b>	<b>41</b>
6.1	Деякі важливі нерівності . . . . .	41
6.2	Види збіжностей . . . . .	41
6.3	Закон великих чисел . . . . .	43
6.3.1	Для незалежних та однаково розподілених випадкових величин . . . . .	43
6.3.2	Для незалежних та різнорозподілених випадкових величин . . . . .	44
6.4	Центральна гранична теорема . . . . .	44
6.4.1	Для незалежних та однаково розподілених випадкових величин . . . . .	44

## Вступ

Ви обрали щойно пдф з теорії ймовірності на **нестрогому** рівні. Більшість доведень певних фактів залишаються, але деякі означення розкрити в повній мірі не вийде. Цей варіант призначений для тих, хто не орієнтується в дисципліні "теорія міри".

# 1 Основи теорії ймовірності

## 1.1 Імовірнісний простір

**Definition 1.1.1** Простором вибірок (або часто називають простором елементарних подій) називають множину  $\Omega$ , яка складається з елементів  $\omega \in \Omega$ , які називаються елементарними подіями.

**Example 1.1.2** Кидаємо один раз монету, простір вибірок  $\Omega$  складається з двох точок:  $\Omega = \{H, T\}$ , де  $H = \text{head}$  (орел) та  $T = \text{tail}$  (решка).

Ми виключаємо патові сценарії, коли монета впала на грань чи монета зникла тощо.

**Example 1.1.3** Кидаємо три рази монету, простір вибірок  $\Omega$  уже буде такою:  
 $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ .

**Example 1.1.4** Серед  $M$  куль обираємо  $n$  разів, після кожного вибору цю кулю повертаємо назад.  $\Omega = \{\omega : \omega = (a_1, \dots, a_n) \mid a_i = 1, \dots, M\}$ , у цьому випадку  $a_i$  – це який м'яч ми взяли на  $i$ -му кроці обирання. Зауважимо, що  $|\Omega| = M^n$ . Вибірка  $(a_1, \dots, a_n)$  в круглих дужках означає, що вона впорядкована. Тобто в цьому випадку порядок важливий.

$\Omega = \{\omega : \omega = [a_1, \dots, a_n] \mid a_i = 1, \dots, M\}$ , у цьому випадку  $a_i$  – це який м'яч ми взяли на  $i$ -му кроці обирання. Зауважимо, що  $|\Omega| = \tilde{C}_M^n$ . Вибірка  $[a_1, \dots, a_n]$  в квадратних дужках означає, що вона неупорядкована. Тобто в цьому випадку порядок неважливий.

**Example 1.1.5** Серед  $M$  куль обираємо  $n \leq M$  разів, після вибору цю кулю НЕ повертаємо назад.  
 $\Omega = \{\omega : \omega = (a_1, \dots, a_n) \mid a_k \neq a_l, k \neq l; a_i = 1, \dots, M\}$ . Зауважимо, що  $|\Omega| = A_M^n$ .  
 $\Omega = \{\omega : \omega = [a_1, \dots, a_n] \mid a_k \neq a_l, k \neq l; a_i = 1, \dots, M\}$ . Зауважимо, що  $|\Omega| = C_M^n$ .

**Example 1.1.6** Опишемо зустріч двох осіб, що узгодили між собою перетнутися протягом однієї години. Позначимо за  $x$  – час приїзду першої особи;  $y$  – час приїзду другої особи. Маємо такий простір елементарних подій:  $\Omega = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$ .

**Definition 1.1.7** Задано  $\Omega$  – простір елементарних подій.

Підмножину  $A \subset \Omega$  будемо називати **випадковою подією**. Множина  $\emptyset \subset \Omega$  називається **неможливою подією (impossible event)**. Множина  $\Omega \subset \Omega$  називається **достовірною подією (certain event)**.

**Example 1.1.8** Кидаємо три рази монету, у нас  $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ . Опишемо подію  $A$ , де хоча б два орли (head) випали. Це буде  $A = \{HHT, HTH, THH, HHH\}$ .

Ми знаємо, що над множинами можна проводити операції:  $\cup$ ,  $\cap$ , взяття доповнення тощо. Якщо переформулювати це в жаргон теорії ймовірностей, то ми отримаємо наступний опис:

$A \cup B$  – подія, яка відбудеться, коли хоча б одна з подій ( $A$  або  $B$ ) відбудеться.

$A \cap B$  – подія, яка відбудеться, коли одночасно ці дві події ( $A$  та  $B$ ) відбудуться.

$A^c$  – подія, яка відбудеться тоді, коли не відбудеться подія  $A$ .

$A \setminus B$  – подія, яка відбудеться, коли спрацює подія  $A$  та не спрацює подія  $B$ .

**Example 1.1.9** Кидаємо двічі монету, маємо простір  $\Omega = \{HH, HT, TH, TT\}$ . Опишемо  $A = \{HH, HT, TH\}$  – випав хоча б один орел;  $B = \{TT, TH, HT\}$  – випала хоча б одна решка.

$A \cup B = \{HH, HT, TH, TT\} = \Omega$  – випала або хоча б одна решка, або випав хоча б один герб (це зрозуміло, що буде достовірною подією).

$A \cap B = \{HT, TH\}$  – випала хоча б одна решка та випав хоча б один герб.

$A^c = \{TT\}$  – не випав жодний орел.

**Definition 1.1.10** Події  $A, B \subset \Omega$  називаються **несумісними**, якщо

$$A \cap B = \emptyset$$

Події  $A_1, A_2, \dots \subset \Omega$  називаються **попарно несумісними**, якщо

$$A_i \cap A_j = \emptyset, i \neq j$$

**Definition 1.1.11** Маємо простір елементарних подій  $\Omega$ . Створимо  $\mathcal{F}$  – клас всіх підмножин  $\Omega$ , що буде задовольняти таким властивостям:

$$\begin{aligned}\Omega &\in \mathcal{F}; \\ A \in \mathcal{F} &\implies A^c \in \mathcal{F}; \\ A_1, A_2, \dots \in \mathcal{F} &\implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}\end{aligned}$$

Такий клас підмножин  $\mathcal{F}$  називають  **$\sigma$ -алгеброю подій**.

Пара  $(\Omega, \mathcal{F})$  називають **вимірним простором**.

**Proposition 1.1.12** Нехай  $\Omega$  – простір елементарних подій та  $\mathcal{F}$  –  $\sigma$ -алгебра подій. Тоді:

- 1)  $\emptyset \in \mathcal{F}$ ;
- 2)  $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$ ;
- 3)  $A, B \in \mathcal{F} \implies A \setminus B \in \mathcal{F}$ .

**Proof.**

Доведення неважке.

- 1) Оскільки  $\Omega \in \mathcal{F}$ , то звідси  $\Omega^c = \emptyset \in \mathcal{F}$ .
- 2) Нехай  $A, B \in \mathcal{F}$ . Тоді  $A \cap B = (A^c \cup B^c)^c \in \mathcal{F}$ .
- 3) Нехай  $A, B \in \mathcal{F}$ . Маємо  $A \setminus B = A \cap B^c$ . За другою властивістю,  $A \setminus B \in \mathcal{F}$ .

Усі властивості доведені. ■

**Remark 1.1.13** Якщо простір елементарних подій  $\Omega$  буде скінченним або зліченим, то  $\sigma$ -алгебру подій можна покласти за  $\mathcal{F} = 2^\Omega$ .

**Definition 1.1.14** Задано  $(\Omega, \mathcal{F})$  – вимірний простір.

**Імовірністю** назвемо функцію множин  $\mathbb{P}: \mathcal{F} \rightarrow \mathbb{R}$ , яка задовольняє трьом аксіомам Колмогорова:

$$\begin{aligned}1) \quad &\forall A \in \mathcal{F} : \mathbb{P}(A) \geq 0; \\ 2) \quad &\forall A_1, A_2, \dots \in \mathcal{F} : \forall i \neq j : A_i \cap A_j = \emptyset : \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n); \\ 3) \quad &\mathbb{P}(\Omega) = 1.\end{aligned}$$

**Remark 1.1.15** Насправді кажучи, перші дві аксіоми для  $\mathbb{P}$  означають, що  $\mathbb{P}$  – це міра.

**Proposition 1.1.16** Властивості ймовірності

Справедливе наступне:

- 1)  $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ ;
- 2)  $\mathbb{P}(\emptyset) = 0$ ;
- 3)  $A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$ ;
- 4)  $\mathbb{P}(A) \leq 1$ ;
- 5)  $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ ;
- 6)  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$  (цю формулу можна узагальнити на довільне скінченне число множин, як це було з потужностями);
- 7)  $\mathbb{P}$  неперервна зверху та знизу одночасно.

Всі ці властивості – це випливає з теорії міри.

## 1.2 Класична модель ймовірності

Припустимо, що  $\Omega$  – злічений простір елементарних подій. Маємо  $\mathbb{P}$  – ймовірність, що задана на  $\sigma$ -алгебрі  $\mathcal{F}$ . У силу зліченності маємо наступне:

$$\mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{\omega_i\}\right) = \sum_{n=1}^{\infty} \mathbb{P}(\{\omega_i\}).$$

Кожному  $\omega_i \in \Omega$  поставимо число  $\mathbb{P}(\{\omega_i\}) = p_i \geq 0$ .

Тоді кожна подія  $A$  має ймовірність  $\mathbb{P}(A) = \sum_{\omega_i \in A} p_i$ .

Всі ці процедури можна аналогічно повторити, якщо  $\Omega$  – скінченний простір.

Якщо  $|\Omega| = k$  – скінченний, то прийнято вважати, що всі елементарні події  $\omega_i$  рівноможливі, тобто всі  $p_1 = p_2 = \dots = p_k \stackrel{\text{позн.}}{=} p$  рівні. Це означатиме, що  $\sum_{n=1}^k p_i = p_1 + \dots + p_k = kp = 1$ . Отже, кожна ймовірність  $p = \frac{1}{k}$ .

Тоді кожна подія  $A$  має ймовірність  $\mathbb{P}(A) = \sum_{\omega_i \in A} p_i = \frac{m}{k}$ , де  $m$  – кількість елементарних подій в  $A$ .

Отже, в класичній моделі ймовірності при  $\Omega$  – скінченна, ймовірність події обчислюється так:

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

**Example 1.2.1** Кидаємо один раз монету,  $\Omega = \{H, T\}$ . Тоді в класичному сенсі  $\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = \frac{1}{2}$ . Тобто ймовірність випадіння орла чи решки – рівноможлива.

**Example 1.2.2** Кидаємо двічі монету. Знайти ймовірність того, що хоча б один орел випаде. Маємо  $\Omega = \{HH, HT, TH, TT\}$ , описана подія  $A = \{HH, HT, TH\}$ . Тоді звідси  $\mathbb{P}(A) = \frac{3}{4}$ .

### Example 1.2.3 Задача про буху секретарку

Секретарка поклала  $n$  листів в  $n$  чистих конвертів, заклеїла ці конверти та після цього написала адреси. Знайти ймовірність того, що хоча б один з листів дійде за призначенням.

Позначимо  $A_i$  – подія, де  $i$ -ий лист дійшов за призначенням. Тоді маємо, що  $\mathbb{P}(A_i) = \frac{1}{n}$ .

Маємо подію  $A$  – подія, де хоча б один з листів дійде за призначенням. Тобто  $A = \bigcup_{i=1}^n A_i$ .

$$\mathbb{P}(A_i) = \frac{1}{n};$$

$$\mathbb{P}(A_i \cap A_j) = \frac{1}{n(n-1)};$$

$$\mathbb{P}(A_i \cap A_j \cap A_k) = \frac{1}{n(n-1)(n-2)};$$

$\vdots$

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \frac{1}{n!}.$$

Звідси випливає, що за формулою включення-виключення,

$$\mathbb{P}(A) = n \cdot \frac{1}{n} - C_n^2 \frac{1}{n(n-1)} + \dots + (-1)^{n-1} \frac{1}{n!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!}.$$

Спрямуємо  $n \rightarrow \infty$ , зауважимо, що  $e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots$ . Тоді звідси  $\lim_{n \rightarrow \infty} \mathbb{P}(A) = 1 - \frac{1}{e} \approx 0.63$ .

## 1.3 Геометрична ймовірність

Нехай  $\Omega \subset \mathbb{R}^n$  – деяка замкнена область, яка буде інтерпретуватися як простір елементарних подій, кожна підмножина  $A \subset \mathbb{R}^n$  – подія. Причому всі елементарні події рівноможливі. У якості  $\sigma$ -алгебри подій  $\mathcal{F}$  беремо ті підмножини, що мають міру  $\mu$ . Тоді в якості ймовірності беремо

$$\mathbb{P}(A) = \frac{\mu(A)}{\mu(\Omega)}$$

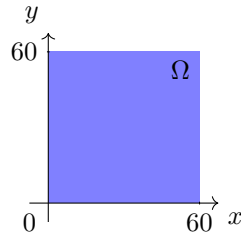
$\mathbb{R}^1$ :  $\mu$  – довжина;

$\mathbb{R}^2$ :  $\mu$  – площа;

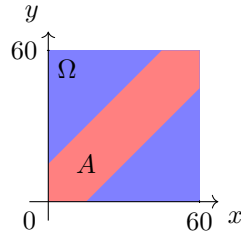
$\mathbb{R}^3$ :  $\mu$  – об'єм.

**Example 1.3.1** Саша й Маша домовилися про зустріч між 10 та 11 годиною. Кожен із них приходить у випадковий момент проміжку. Якщо хтось чекатиме протягом 15 хвилин, людина піде зі зустрічі. Знайти ймовірність того, що зустріч відбудеться.

Маємо  $\Omega = \{(x, y) \mid x, y \in [0, 60]\}$ , де  $x$  – момент приходу Саші та  $y$  – момент приходу Маші (якщо, наприклад,  $x = 30$ , то Саша прийшла о 10:30). Даний простір елементарних подій можна намалювати геометрично ось так:



Позначимо  $A$  – подія, яка описує, що зустріч відбудеться. Тобто  $A = \{(x, y) \in \Omega : |x - y| \leq 15\}$ . Під виразом  $|x - y|$  мається на увазі тривалість приходу між Сашею та Машею.



Нас цікавить ймовірність потрапити в червону область.

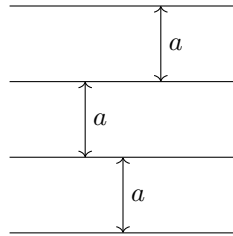
Оскільки геометрично ми в двовимірному випадку, то  $\mathbb{P}(A) = \frac{S(A)}{S(\Omega)}$ . Обчислимо площу кожної

фігури:  $S(\Omega) = 60^2 = 3600$ ;  $S(A) = 60^2 - 2 \cdot \frac{45^2}{2} = 1575$ .

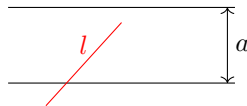
Таким чином,  $\mathbb{P}(A) = \frac{1575}{3600} = \frac{7}{16}$ .

### Example 1.3.2 Задача Бюффона

Намалюємо на площині паралельні прямі, вони всі на відстані  $a$  між собою. На цю площину ми кидаємо голку довжини  $l$ , причому  $l < a$ . Знайти ймовірність того, що голка перетне будь-яку одну пряму.



За умовою довжини голки, нам буде досить розглядати дві прямі.

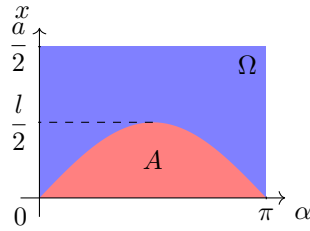


Положення голки Бюффон задає парами  $(\alpha, x)$ , де  $\alpha \in [0, \pi]$  – кут між прямою та голкою, а також  $x \in [0, \frac{a}{2}]$  – відстань від центра голки до найближчої прямої. У такому разі простір елементарних подій  $\Omega = \{(\alpha, x) \in \mathbb{R}^2 \mid \alpha \in [0, \pi], x \in [0, \frac{a}{2}]\}$ .

Позначимо  $A$  – подія, коли голка перетинає пряму. Описується ось так:  $A = \left\{(\alpha, x) \in \Omega : x < \frac{l}{2} \sin \alpha\right\}$ .

Тепер намалюємо ці дві множини на площині





Знову цікавить ймовірність потрапити в червону область.

$$\text{У цьому випадку } \mathbb{P}(A) = \frac{S(A)}{S(\Omega)} = \frac{\int_0^\pi \frac{l}{2} \sin \alpha \, d\alpha}{\pi \cdot \frac{a}{2}} = \dots = \frac{2l}{a\pi}.$$

На цьому не все. Ця задача використовується для стохастичного обчислення чотирьох та суботусяла  $\pi$ . Зокрема із рівності  $\mathbb{P}(A) = \frac{2l}{a\pi}$  випливатиме, що  $\pi = \frac{2l}{a\mathbb{P}(A)}$ . Припустимо, що  $\frac{l}{a} = 0.8$ . Якщо кинути голку 5000 разів, серед яких 2532 разів перетне пряму, то приблизно  $\mathbb{P}(A) \approx 0.5064$ . Значить, тоді  $\pi \approx 3.1596$ .

**Remark 1.3.3** Геометрична ймовірність, на жаль, заплутана модель для описання експерименту. Може таке бути, що ми описуємо ймовірність однієї події різними моделями, а результат – різний. Відомим прикладом до цього є [парадокс Бертрана](#).

Якщо коротко, маємо коло радіуса  $R$  та вписаний рівносторонній трикутник. Задача була знайти ймовірність того, що навмання обрана хорда матиме довжину меншу за сторону трикутника. Бертран отримав три результати на основі різної моделі:  $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ . Результат різний, оскільки "навмання обрана хорда" там трактувалась по-різному.

## 1.4 Умовна ймовірність

**Definition 1.4.1** Маємо  $(\Omega, \mathcal{F}, \mathbb{P})$  – ймовірнісний простір. Припустимо, що подія  $B$  має ймовірність  $\mathbb{P}(B) > 0$ .

Ймовірність події  $A$  за умовою, що спрацювала подія  $B$ , називається **умовною ймовірністю  $A$  за умовою  $B$** . Обчислюється за такою формулою:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Щоб чітко зрозуміти, чому таке означення, ми розглянемо на прикладі на класичній моделі.

**Example 1.4.2** Підкидаємо монету двічі. Ми зараз ілюстративно розв'яжемо дві задачі:

I. Знайти ймовірність того, що випаде хоча б одна решка.

У нас  $\Omega = \{HH, HT, TH, TT\}$ , наша подія  $A = \{HT, TH, TT\}$ . Ілюстративно це буде виглядати таким чином:

$\Omega$		$\Omega$	
HT	TT	HT	TT
HH	TH	HH	TH

Червона – подія  $A$ . Синій квадрат – подія  $\Omega$ .

У класичній імовірності ми робили  $\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{4}$ .

II. Знайти ймовірність того, що випаде хоча б одна решка за умовою, що на першому випав орел.

На відміну від першого випадку, нас більше не цікавить випадок  $TT$  та  $TH$ . Всі можливі події залишимо синім кольором, а необхідне червоним.

$\Omega$		$\Omega$	
$HT$	$TT$	$HT$	$TT$
$HH$	$TH$	$HH$	$TH$

Червона – подія  $A|B$ . Синій квадрат – подія  $B$ .

У нас тепер достовірна подія  $B = \Omega \cap B = \{HT, HH\}$ . Ось це "за умовою" вже звужує наш простір елементарних подій. Сама подія  $A|B = \{HT\} = A \cap B$ . Отже,  $\mathbb{P}(A|B) = \frac{|A \cap B|}{|B|} = \frac{1}{2}$ . Із іншого

боку, ми можемо цю формулу переписати ось так:  $\mathbb{P}(A|B) = \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$  (такі

міркування працюють лише в класичній моделі). Для інших моделей ми хочемо, щоб ця формула теж працювала.

**Example 1.4.3** Кидаємо дві гральні кісточки d6. Ми хочемо знайти ймовірність того, що випаде більше шести очок за умовою, що перша кістка повернула три очки.

Маємо простір  $\Omega = \{\omega : \omega = (a, b) : 1 \leq a \leq 6, 1 \leq b \leq 6\}$ .

$A$  – подія, яка описує, що ми отримали більше шести очок (наприклад,  $(4, 6) \in A$ );

$B$  – подія, яка описує, що перша кістка дає три очки (наприклад,  $(3, 5) \in B$ ).

Тоді  $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ .

Зауважимо, що  $\mathbb{P}(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{3}{36}$ . У нас всього три події можливі:  $(3, 4), (3, 5), (3, 6)$ .

Також маємо  $\mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{6}{36}$ . Тут можливі шість подій:  $(3, 1), (3, 2), \dots, (3, 6)$ .

Остаточно отримаємо ймовірність  $\mathbb{P}(A|B) = \frac{3}{6} = \frac{1}{2}$ .

Для порівняння, ймовірність того, що випаде більше шести очок (без жодних інших умов)  $\mathbb{P}(A) = \frac{21}{36} = \frac{7}{12}$  – вийшла більша ймовірність.

#### Proposition 1.4.4 Властивості умовної ймовірності

Справедливе наступне:

- 1)  $\mathbb{P}(A|B) \geq 0$ ;
- 2)  $\mathbb{P}(B|B) = 1$ ;
- 3)  $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n|B\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n|B)$ ;
- 4)  $\mathbb{P}(\emptyset|A) = 0$ ;
- 5)  $\mathbb{P}(A|B) = 1$ , якщо  $A \supset B$ .

Вправа: довести.

**Corollary 1.4.5** Трійка  $(\Omega \cap B, \mathcal{F} \cap B, \mathbb{P}(\cdot|B))$  задає ймовірнісний простір.

## 1.5 Незалежність подій

**Definition 1.5.1** Події  $A, B$  називаються **незалежними**, якщо

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

#### Proposition 1.5.2 Властивості незалежних подій

Нехай  $A, B$  – незалежні події. Справедливе наступне:

- 1)  $\mathbb{P}(A|B) = \mathbb{P}(A)$        $\mathbb{P}(B|A) = \mathbb{P}(B)$ ;

- 2)  $A, B^c$  або  $A^c, B$  або  $A^c, B^c$  – також незалежні події;  
 3) Нехай  $A, B$  та  $A, C$  між собою незалежні та  $B, C$  несумісні. Тоді  $A, B \sqcup C$  – незалежні.  
*Вправа: довести.*

**Definition 1.5.3** Події  $A_1, \dots, A_n$  називаються **незалежними (у сукупності)**, якщо

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

**Definition 1.5.4** Події  $A_1, \dots, A_n$  називаються **попарно незалежними**, якщо

$$\forall i, j : i \neq j : \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j)$$

**Remark 1.5.5** Із того, що  $A_1, \dots, A_n$  незалежні у сукупності, випливає, що  $A_1, \dots, A_n$  – попарно незалежні.

Проте навпаки не працює.

### Example 1.5.6 Приклад Бернштейна

Маємо тетраедр, три грані якого розфарбовані відповідно червоним, зеленим та синім кольорами. Четверта грань – комбінація цих трьох кольорів.

$R$  – подія, коли тетраедр показує червоний колір після підкидання;

$G$  – подія, коли тетраедр показує зелений колір після підкидання;

$B$  – подія, коли тетраедр показує синій колір після підкидання.

Зауважимо спочатку, що  $\mathbb{P}(R) = \mathbb{P}(G) = \mathbb{P}(B) = \frac{1}{2}$ .

Події  $R, G, B$  попарно незалежні. Дійсно,  $\mathbb{P}(R \cap G) = \mathbb{P}(G \cap B) = \mathbb{P}(B \cap R) = \frac{1}{4}$ .

Події  $R, G, B$  не є незалежними у сукупності. Дійсно,  $\mathbb{P}(R \cap G \cap B) = \frac{1}{4} \neq \frac{1}{8}$ .

### Theorem 1.5.7 Теорема множення ймовірностей

$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|(A_1 \cap A_2)) \dots \mathbb{P}(A_n|(A_1 \cap A_2 \cap \dots \cap A_{n-1}))$ .

#### Proof.

Це можна довести методом МІ. Тільки напишу базу індукції, а далі буде неважко довести. Ми знаємо, що  $\mathbb{P}(A_2|A_1) = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)}$  за означенням. Звідси випливатиме формула  $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1)$ . ■

## 1.6 Формула повної ймовірності

**Definition 1.6.1** Заданий  $\Omega$  – простір елементарних подій.

Набір подій  $\{A_1, A_2, \dots\}$  утворюють **повну групу подій**, якщо

$$\forall i \neq j : A_i \cap A_j = \emptyset$$

$$\bigcup_{n=1}^{\infty} A_n = \Omega$$

Маємо  $\Omega$  – простір елементарних подій. Припустимо, що  $\{H_1, H_2, \dots, H_n\}$  утворюють повну групу подій, у такому разі такі події називають **припущеннями**.

Нехай подія  $A$  відбувається разом з якимось припущенням. Тоді

$$\mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \mathbb{P}(A \cap (H_1 \sqcup \dots \sqcup H_n)) = \mathbb{P}((A \cap H_1) \sqcup \dots \sqcup (A \cap H_n)) = \sum_{k=1}^n \mathbb{P}(A \cap H_k) = \sum_{k=1}^n \mathbb{P}(H_k) \mathbb{P}(A|H_k).$$

Таким чином, ми отримали формулу повної ймовірності:

$$\mathbb{P}(A) = \sum_{k=1}^n \mathbb{P}(H_k) \mathbb{P}(A|H_k)$$

**Example 1.6.2** У магазин постачають 80% телефонів з Китаю, 15% з В'єтнаму, 5% з Південної Кореї. Із кожної країни кількість бракованих відповідно 1%, 0.1%, 0.01%. Обчислити ймовірність того, що куплений телефон буде бракованим.

$H_1$  – подія, де телефон із Китаю;

$H_2$  – подія, де телефон із В'єтнаму;

$H_3$  – подія, де телефон із Південної Кореї.

Позначимо  $A$  – подія, де куплений телефон бракований (тут не зазначається, із якої країни ми купляємо телефон). Оскільки  $H_1, H_2, H_3$  утворюють повну групу подій, то звідси  $\mathbb{P}(A) = \mathbb{P}(H_1)\mathbb{P}(A|H_1) + \mathbb{P}(H_2)\mathbb{P}(A|H_2) + \mathbb{P}(H_3)\mathbb{P}(A|H_3)$ .

Зауважимо, що  $\mathbb{P}(H_1) = 0.8$ ,  $\mathbb{P}(H_2) = 0.15$ ,  $\mathbb{P}(H_3) = 0.05$  (за умовою).

Також зазначимо, що  $\mathbb{P}(A|H_1) = 0.01$ ,  $\mathbb{P}(A|H_2) = 0.001$ ,  $\mathbb{P}(A|H_3) = 0.0001$ .

Отже, підставивши всі значення, отримаємо  $\mathbb{P}(A) = 0.008155$ .

## 1.7 Формула Баєса

Маємо  $\Omega$  – простір елементарних подій. Припустимо, що  $\{H_1, H_2, \dots, H_n\}$  утворюють повну групу подій, у такому разі такі події називають **припущеннями**. У результаті проведення експерименту відбулася подія  $A$ .

Виникає питання: чому дорівнює  $\mathbb{P}(H_1|A), \dots, \mathbb{P}(H_n|A)$ ?

$$\mathbb{P}(H_i|A) = \frac{\mathbb{P}(H_i \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(H_i)\mathbb{P}(A|H_i)}{\sum_{k=1}^n \mathbb{P}(H_k)\mathbb{P}(A|H_k)}.$$

Таким чином, отримали формулу Баєса:

$$\mathbb{P}(H_i|A) = \frac{\mathbb{P}(H_i)\mathbb{P}(A|H_i)}{\sum_{k=1}^n \mathbb{P}(H_k)\mathbb{P}(A|H_k)}$$

Така формула дає відповідь на питання, чому дорівнює ймовірність того, що мала місце кожна з гіпотез  $H_i$ , якщо вже подія  $A$  сталася.

До речі,  $\mathbb{P}(H_i)$  називають часто **апостеріорною ймовірністю**, а також  $\mathbb{P}(H_i|A)$  називають **апостеріорною ймовірністю**.

**Example 1.7.1** Всього 10 студентів у групі, серед них: троє вчаться на "5" (відмінники), четверо – на "4" (хорошисти), двоє – на "3" (середняки) та один – на "2" (двієчники). Викладач підготував 20 екзаменаційних питань. Відмінники знають всі питання, хорошисти – 16, середняки – 10, а двієчник – лише 5. Білет складається з 3 питань. Деякий студент відповів на всі 3 питання. Знайти ймовірність того, що він – двієчник.

$H_1$  – подія, де студент – відмінник;

$H_2$  – подія, де студент – хорошист;

$H_3$  – подія, де студент – середняк;

$H_4$  – подія, де студент – двієчник.

Зрозуміло, що  $\mathbb{P}(H_1) = \frac{3}{10}$ ,  $\mathbb{P}(H_2) = \frac{4}{10}$ ,  $\mathbb{P}(H_3) = \frac{2}{10}$ ,  $\mathbb{P}(H_4) = \frac{1}{10}$ .

Позначимо  $A$  – подія, де студент відповів на всі три питання. За формулою повної ймовірності,  $\mathbb{P}(A) = \mathbb{P}(H_1)\mathbb{P}(A|H_1) + \mathbb{P}(H_2)\mathbb{P}(A|H_2) + \mathbb{P}(H_3)\mathbb{P}(A|H_3) + \mathbb{P}(H_4)\mathbb{P}(A|H_4)$ .

Зазначимо, що  $\mathbb{P}(A|H_1) = 1$ ,  $\mathbb{P}(A|H_2) = \frac{16 \cdot 15 \cdot 14}{20 \cdot 19 \cdot 18}$ ,  $\mathbb{P}(A|H_3) = \frac{10 \cdot 9 \cdot 8}{20 \cdot 19 \cdot 18}$ ,  $\mathbb{P}(A|H_4) = \frac{5 \cdot 4 \cdot 3}{20 \cdot 19 \cdot 18}$ .

Отже, звідси  $\mathbb{P}(A) = \frac{197}{380}$ . Наша задача: знайти  $\mathbb{P}(H_4|A)$ . Тоді за формулою Баєса,

$$\mathbb{P}(H_4|A) = \frac{\mathbb{P}(H_4) \cdot \mathbb{P}(A|H_4)}{\mathbb{P}(A)} = \frac{1}{591}.$$

## 1.8 Схема Бернуллі

Маємо експеримент із  $n$  незалежними випробуваннями, в кожному з них успіх або невдача. Позначимо  $\mathbb{P}\{\text{успіх}\} = p$ .

Задача: знайти ймовірність того, що із  $n$  випробувань рівно  $k$  разів будуть успішними. Таку ймовірність часто позначають за  $p_n(k)$ .

Спочатку треба з'ясувати, скільки всього існують варіантів отримати  $k$  успіхів (один із можливих варіантів – це коли перші  $k$  випробувань поспіль успішні, а решта – ні; нам потрібні всі можливості). Перефразуємо в комбінаторну задачу: у  $n$  комірок розташувати  $k$  успіхів. Оскільки порядок неважливий, то всього в нас  $C_n^k$  таких способів.



Зелені кола – це  $k$  успіхів та решта червоні кола – це  $n - k$  невдач.

Імовірність отримати  $k$  успіхів та  $n - k$  невдач (в одному із способів) дорівнює  $p^k(1 - p)^{n-k}$  в силу незалежності випробувань.

Разом ми отримаємо нашу бажану ймовірність:

$$p_n(k) = C_n^k p^k (1 - p)^{n-k}$$

Часто ймовірність невдачі, тобто  $1 - p$ , позначають за  $q$  для зручності. Тобто  $q = 1 - p$ .

**Example 1.8.1** Розглянемо експеримент, коли ми тричі підкидуємо монету. Знайдемо ймовірності того, що серед 3 підкидувань випаде: 0 орлів, 1 орел, 2 орли, 3 орли.

Зазначимо, що  $p = \mathbb{P}\{H\} = \frac{1}{2}$ , де подія  $\{H\}$  – випав орел – наш успіх.

$$p_3(0) = C_3^0 \left(\frac{1}{2}\right)^0 \left(1 - \frac{1}{2}\right)^{3-0} = \frac{1}{8};$$

$$p_3(1) = C_3^1 \left(\frac{1}{2}\right)^1 \left(1 - \frac{1}{2}\right)^{3-1} = \frac{3}{8};$$

$$p_3(2) = C_3^2 \left(\frac{1}{2}\right)^2 \left(1 - \frac{1}{2}\right)^{3-2} = \frac{3}{8};$$

$$p_3(3) = C_3^3 \left(\frac{1}{2}\right)^3 \left(1 - \frac{1}{2}\right)^{3-3} = \frac{1}{8}.$$

Рахували ми стільки ймовірностей, щоб помітити таку ситуацію. Найбільша ймовірність становить  $\frac{3}{8}$ . Отже, в даному експерименті найімовірніше за всього будуть або 1 успіх, або 2 успіхи.

**Example 1.8.2** Розглянемо експеримент, коли ми уже чотири рази підкидуємо монету. Знайдемо ймовірності того, що серед 4 підкидувань випаде: 0 орлів, 1 орел, 2 орли, 3 орли, 4 орли.

Я тепер детально це розписувати не буду, просто запишу відповідні значення:

$$p_4(0) = \frac{1}{16}, \quad p_4(1) = \frac{4}{16}, \quad p_4(2) = \frac{6}{16}, \quad p_4(3) = \frac{4}{16}, \quad p_4(4) = \frac{1}{16}.$$

Знову найбільша ймовірність тут  $\frac{6}{16}$ . Отже, в даному експерименті найімовірніше за всього будуть 2 успіхи.

Це лише найпростіші задачі, де можна ручками знайти найімовірнішу кількість успіхів. Але все одно хочеться знайти цю кількість, не рахуючи всі ці ймовірності.

## 1.9 Найімовірніша кількість успіхів у схемі Бернуллі

Позначимо  $k^* = \arg\max_{k=0, \dots, n} p_n(k)$  – тобто таке значення аргументу  $k$ , при якому досягається найбільше значення ймовірності.

Наша задача: знайти  $k^*$  – найімовірніша кількість успіхів у схемі Бернуллі. Маємо таку систему

$$\text{нерівностей: } \begin{cases} p_n(k^*) \geq p_n(k^* - 1) \\ p_n(k^*) \geq p_n(k^* + 1) \end{cases}.$$

Розпишемо еквівалентні перетворення:

$$\begin{aligned} \begin{cases} p_n(k^*) \geq p_n(k^* - 1) \\ p_n(k^*) \geq p_n(k^* + 1) \end{cases} &\iff \begin{cases} \frac{n!}{k^*!(n - k^*)!} p^{k^*} q^{n - k^*} \geq \frac{n!}{(k^* - 1)!(n - (k^* - 1))!} p^{k^* - 1} q^{n - (k^* - 1)} \\ \frac{n!}{k^*!(n - k^*)!} p^{k^*} q^{n - k^*} \geq \frac{n!}{(k^* + 1)!(n - (k^* + 1))!} p^{k^* + 1} q^{n - (k^* + 1)} \end{cases} \iff \\ &\iff \begin{cases} \frac{p}{k^*} \geq \frac{q}{n - (k^* - 1)} \\ \frac{q}{n - k^*} \geq \frac{p}{k^* + 1} \end{cases} \iff \begin{cases} k^* \leq np + p \\ k^* \geq np - q \end{cases} \iff \begin{cases} k^* \leq np + p \\ k^* \geq np + p - 1 \end{cases}. \end{aligned}$$

Отже, можемо зробити висновок:

I. Якщо  $np + p \in \mathbb{Z}$ , то тоді існуватимуть два числа:  $k_1^* = np - q$  або  $k_2^* = np + p$ .

II. Якщо  $np + p \notin \mathbb{Z}$ , то тоді число  $k^* = [np + p]$ .

**Remark 1.9.1** Під час знаходження  $k^*$  ми довели, що  $p_n(k)$  спочатку зростає за  $k$  до певного числа, потім спадає. Максимальне значення досягається завжди.

**Example 1.9.2** Ми 100 разів підкидаємо кістку d6. Успіхом будемо вважати той випадок, коли випало число 6. З'ясуємо, яка найімовірніша кількість успіху.

$n = 100$  та  $p = \frac{1}{6}$  в нашому випадку. Зазначимо, що  $np + p = \frac{100}{6} + \frac{1}{6} = \frac{101}{6} \notin \mathbb{Z}$ . Таким чином,  $k^* = [np + p] = 16$ .

**Example 1.9.3** Повернемось до задачі, де ми 3 рази кидаємо монету та вважаємо випадіння герба за успіхом. Іншим способом знайдемо найімовірнішу кількість успіхів.

$n = 3$  та  $p = \frac{1}{2}$  в нашому випадку. Зазначимо, що  $np + p = 2 \in \mathbb{Z}$ . Таким чином,  $k_1^* = np + p = 2$  або  $k_2^* = np + p - 1 = 1$ .

## 1.10 Поліноміальна схема

Нехай  $A_1, \dots, A_m$  утворюють повну групу подій, позначимо їхні ймовірності  $\mathbb{P}(A_i) = p_i$ .

Проводимо  $n$  разів незалежні випробування, в кожному з яких відбувається одна з подій  $A_i$ .

Задача: знайти ймовірність того, що:

подія  $A_1$  виникне  $k_1$  разів;

подія  $A_2$  виникне  $k_2$  разів;

$\vdots$

подія  $A_m$  виникне  $k_m$  разів.

Зрозуміло, що тут  $k_1 + k_2 + \dots + k_m = n$ . Також допускається, що одна з подій може не відбутися.

Таку ймовірність позначу за  $p_n(k_1, k_2, \dots, k_m)$ .

Якщо згадати комбінаторику, то аналогічним чином (як ми працювали зі схемою Бернуллі) ми отримаємо таку формулу:

$$p_n(k_1, \dots, k_m) = C_n^{k_1, k_2, \dots, k_m} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}$$

Така ймовірність називається поліноміальною.

**Remark 1.10.1** При  $n = 2$  ми повертаємось до схеми Бернуллі. Тут  $A_1$  – успіх та  $A_2 = A_1^c$  – невдача, а також  $p_1 = p$ ,  $p_2 = 1 - p = q$ .

## 2 Випадкові величини

### 2.1 Перші означення

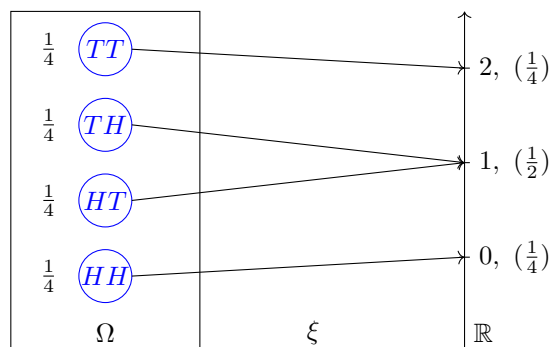
**Definition 2.1.1** Заданий  $(\Omega, \mathcal{F})$  – вимірний простір.

Функцію  $\xi: \Omega \rightarrow \mathbb{R}$  називають **випадковою величиною**. Тобто ця функція описує числовий результат експерименту.

*Увага! Дане означення не є строгим. Про це згодом.*

**Example 2.1.2** Розглянемо модель підкидання монети двічі. Визначимо випадкову величину  $\xi$  – кількість решок (tail). Нагадаю, що  $\Omega = \{HH, HT, TH, TT\}$ . У такому разі випадкова величина діє таким чином:

$$\xi(HH) = 0 \quad \xi(HT) = \xi(TH) = 1 \quad \xi(TT) = 2.$$



Схематичне представлення функції розподілу. Кожній елементарній події ставиться ймовірність. Кожному числу кількості решок за два підкидання також ставиться своя ймовірність.

Нас цікавлять випадкові величини, оскільки ми зможемо таким чином рахувати (скажімо)  $\mathbb{P}\{\xi \in [a, b]\}$ . Але для цього нам потрібно, щоб в нашу  $\sigma$ -алгебру подій  $\mathcal{F}$  потрапляли  $(a, b)$ ,  $(-\infty, a)$ ,  $[a, b]$  тощо. При незліченній множині  $\Omega$  ми не зможемо оголосити  $\mathcal{F} = 2^\Omega$ , інакше це призведе до деяких парадоксів.

Виявляється, що нам буде досить вимагати для випадкових величин, щоб  $\{\xi \in (-\infty, x]\} \in \mathcal{F}$  для всіх  $x \in \mathbb{R}$ . Таким чином, ми можемо отримати  $(a, b)$ ,  $[a, b]$ ,  $(a, b]$ ,  $[a, b)$ ,  $(-\infty, a)$ ,  $[b, +\infty)$ ,  $(b, +\infty) \in \mathcal{F}$ .

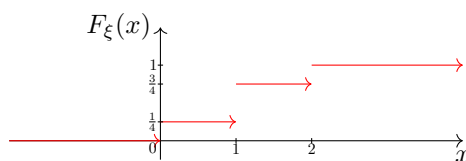
**Definition 2.1.3** Функцією розподілу випадкової величини  $\xi$  називають таку функцію:

$$F_\xi(x) = \mathbb{P}\{\xi \leq x\}, \quad x \in \mathbb{R}$$

**Example 2.1.4** Повернімося до підкидання монети двічі. Мали  $\xi$  – випадкова величина, що позначає кількість решок. Маючи всі дані вище, отримаємо таку функцію розподілу:

$$F_\xi(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}.$$

Наприклад, чому при  $x = 1.5$  ми маємо  $F_\xi(1.5) = \frac{3}{4}$ . Інтуїтивно кажучи, тут написана ймовірність того, що ми отримаємо не більше 1.5 решок (тобто або одна решка, або жодну). Тобто  $\mathbb{P}\{\xi \leq 1.5\} = \mathbb{P}\{\xi = 1\} + \mathbb{P}\{\xi = 0\} = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$ .



Графік функції розподілу випадкової величини.

**Proposition 2.1.5** Властивості функції розподілу

Справедливе наступне:

- 1)  $F_\xi \in [0, 1]$ ;
- 2)  $F_\xi$  – неспадна функція;
- 3)  $\mathbb{P}\{\xi \in (x, y]\} = F_\xi(y) - F_\xi(x)$ ;
- 4)  $F_\xi(x) \rightarrow 0$  при  $x \rightarrow -\infty$ ;  $F_\xi(x) \rightarrow 1$  при  $x \rightarrow +\infty$ ;
- 5)  $F_\xi \in C_{\text{right}}(\mathbb{R})$ ;
- 6)  $F_\xi(x-0) = \mathbb{P}\{\xi < x\}$ .

**Proof.**

Доведемо кожну властивість окремо:

1) Випливає з того, які значення приймає ймовірність.

2) Оберемо  $x_1, x_2 \in \mathbb{R}$  такі, що  $x_2 > x_1$ . Тоді

$$F_\xi(x_2) = \mathbb{P}\{\xi \leq x_2\} = \mathbb{P}\{\{\xi \leq x_1\} \cup \{x_1 < \xi \leq x_2\}\} = \mathbb{P}\{\xi \leq x_1\} + \mathbb{P}\{x_1 < \xi \leq x_2\} \geq F_\xi(x_1).$$

3)  $\mathbb{P}\{\xi \in (x, y]\} = \mathbb{P}\{\xi \leq y\} - \mathbb{P}\{\xi \leq x\} = F_\xi(y) - F_\xi(x)$ .

4) Розглянемо послідовність  $\{x_n, n \geq 1\}$  така, що  $x_n \rightarrow -\infty$ , причому монотонно спадна (\*). Зауважимо, що виконується такий ланцюг:

$$\{\xi \leq x_1\} \supset \{\xi \leq x_2\} \supset \dots$$

Отже, за теоремою про неперервність ймовірності, отримаємо наступне:

$$\lim_{n \rightarrow \infty} F_\xi(x_n) = \lim_{n \rightarrow \infty} \mathbb{P}\{\xi \leq x_n\} = \mathbb{P}\left\{\bigcap_{n=1}^{\infty} \{\xi \leq x_n\}\right\} = \mathbb{P}\{\emptyset\} = 0.$$

Випадок, коли  $F_\xi(x) \rightarrow 1$  при  $x \rightarrow +\infty$ , цілком аналогічний.

5) Нам треба довести, що  $\lim_{x \rightarrow x_0+0} F_\xi(x) = F_\xi(x_0)$ .

Оберемо послідовність  $\{x_n, n \geq 1\}$  так, щоб  $x_n \rightarrow x_0$  праворуч, причому монотонно спадає (\*). Тоді знову ланцюг:

$$\{\xi \leq x_1\} \supset \{\xi \leq x_2\} \supset \dots$$

Отже, за теоремою про неперервність ймовірності, отримаємо наступне:

$$\lim_{n \rightarrow \infty} F_\xi(x_n) = \lim_{n \rightarrow \infty} \mathbb{P}\{\xi \leq x_n\} = \mathbb{P}\left\{\bigcap_{n=1}^{\infty} \{\xi \leq x_n\}\right\} = \mathbb{P}\{\xi \leq x_0\} = F_\xi(x_0).$$

6) Оберемо послідовність  $\{x_n, n \geq 1\}$  така, що  $x_n \rightarrow x_0$  ліворуч, причому монотонно зростає (\*). Зауважимо, що виконується такий ланцюг:

$$\{\xi \leq x_1\} \subset \{\xi \leq x_2\} \subset \dots$$

Отже, за теоремою про неперервність ймовірності, отримаємо наступне:

$$\lim_{n \rightarrow \infty} F_\xi(x_n) = \lim_{n \rightarrow \infty} \mathbb{P}\{\xi \leq x_n\} = \mathbb{P}\left\{\bigcup_{n=1}^{\infty} \{\xi \leq x_n\}\right\} = \mathbb{P}\{\xi < x_0\}.$$

Всі властивості доведені. ■

**Remark 2.1.6** Мабуть, я окремо прокоментую (\*), чому нам досить саме спадної послідовності.

Розглянемо довільну  $\{x_n, n \geq 1\}$  таку, що  $x_n \rightarrow -\infty$ . Простіше від супротивного.

!Припустимо, що  $F_\xi(x_n) \not\rightarrow 0$  при  $n \rightarrow \infty$ , тоді існує  $\varepsilon^* > 0$  такий, що  $\forall N : \exists n \geq N : F_\xi(x_n) \geq \varepsilon^*$ .

Тоді звідси можна відокремити підпослідовність  $\{x_{n_k}, k \geq 1\}$  так, щоб  $x_{n_k} \rightarrow -\infty$ , причому можна спадну взяти (бо  $x_n \rightarrow -\infty$ ). Для спадних послідовностей уже знаємо, що  $F_\xi(x_{n_k}) \rightarrow 0$ . Але тоді

$\exists K : \forall k \geq K : F_\xi(x_{n_k}) < \varepsilon^*$  – суперечність!

Буде все аналогічно

**Remark 2.1.7** Особиве зауваження. Деякі досі визначають функцію розподілу таким чином:

$$F_\xi(x) = \mathbb{P}\{\xi < x\}, \quad x \in \mathbb{R}$$

Тобто тут нерівність тепер строга. Глобально нічого не зміниться, окрім деяких властивостей. Наприклад, в цьому випадку функція  $F_\xi \in C_{\text{left}}(\mathbb{R})$ , але вже не справа.



## 2.2 Основні види випадкових величин

### 2.2.1 Дискретна випадкова величина

**Definition 2.2.1** Випадкова величина  $\xi$  називається **дискретною**, якщо

$\xi$  набуває скінченну чи зліченну кількість значень.

Для того, щоб задати дискретну випадкову величину, крім значень випадкової величини ще треба мати ймовірності, із якими ці значення приймаються.

$$p_i = P\{\xi^{-1}\{x_i\}\} = \mathbb{P}(\xi = x_i), \quad \sum_{i=1}^{\infty} p_i = 1.$$

Часто всі дані зручно записувати в **ряд розподілу**, що виглядає таким чином:

$\xi$	$x_1$	$x_2$	$\dots$	$x_n$	$\dots$
$\mathbb{P}$	$p_1$	$p_2$	$\dots$	$p_n$	$\dots$

Подивимось, як виглядає функція розподілу в такому разі:

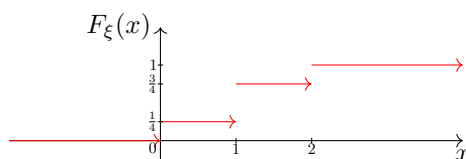
$$F_{\xi}(x) = \begin{cases} 0, & x < x_1 \\ p_1, & x_1 \leq x < x_2 \\ p_1 + p_2, & x_2 \leq x < x_3 \\ \vdots & \\ p_1 + p_2 + \dots = 1, & x \geq x_n \end{cases}$$

Графік даної функції розподілу буде просто ступінчата функція.

**Example 2.2.2** Зокрема в нас був експеримент підкидання монети два рази. Встановлювали випадкову величину  $\xi$  – кількість решок. Це – приклад дискретної випадкової величини, оскільки приймає три значення. Кожна з яких має власну ймовірність. Наведу ряд розподілу:

$\xi$	0	1	2
$\mathbb{P}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Якщо просумувати всі ймовірності, то отримаємо одиницю – тобто це справді ряд розподілу. Як вже було раніше зазначено, ми отримаємо сходишкову функцію розподілу. Не лінь ще раз намалювати.



### 2.2.2 Абсолютно неперервна випадкова величина

**Definition 2.2.3** Випадкова величина  $\xi$  називається **неперервною**, якщо

$$F_{\xi} \in C(\mathbb{R})$$

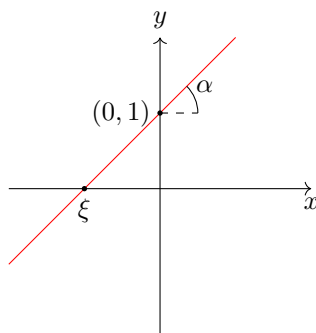
**Proposition 2.2.4**  $\xi$  – неперервна випадкова величина  $\iff \forall x \in \mathbb{R} : \mathbb{P}\{\xi = x\} = 0$ .

**Proof.**

$\Rightarrow$  Дано:  $\xi$  – неперервна випадкова величина, тобто  $F_{\xi} \in C(\mathbb{R})$ . Тоді звідси  $\mathbb{P}\{\xi = x\} = \mathbb{P}\{\xi \leq x\} - \mathbb{P}\{\xi < x\} = F_{\xi}(x) - F_{\xi}(x-0) = 0$ .

$\Leftarrow$  Дано:  $\forall x \in \mathbb{R} : \mathbb{P}\{\xi = x\} = 0$ . Іншими словами,  $\mathbb{P}\{\xi = x\} = \mathbb{P}\{\xi \leq x\} - \mathbb{P}\{\xi < x\} = 0$ , або ще іншими словами,  $F_{\xi}(x) - F_{\xi}(x-0) = 0$ . У силу неперервності справа отримаємо, що  $F_{\xi}(x+0) = F_{\xi}(x-0)$ , а тому функція неперервна в точці  $x$ , зокрема  $F_{\xi} \in C(\mathbb{R})$ . ■

**Example 2.2.5** Розглянемо координатну площину. Через точку  $(0, 1)$  під випадковим кутом  $\alpha \in [0, \pi)$  (подія, що проведемо з таким кутом) проводимо пряму. Необхідно знайти точку перетину з  $OX$ . Згодом описати функцію розподілу.



Позначимо  $\xi$  – випадкова величина, що описує точку перетину з  $OX$ . Зауважимо, що  $\text{ctg } \alpha = \frac{-\xi}{1}$ , тобто звідси  $\xi = -\text{ctg } \alpha$ . Тепер обчислимо функцію розподілу.

$$\begin{aligned} F_{\xi}(x) &= \mathbb{P}\{\xi \leq x\} = \mathbb{P}\{-\text{ctg } \alpha \leq x\} = \mathbb{P}\{\text{ctg } \alpha \geq x\} = \mathbb{P}\{\alpha \in [0, \text{arctg}(-x)]\} = \\ &= \mathbb{P}\{\alpha \in [0, \pi - \text{arctg } x]\} = \frac{\pi - \text{arctg } x}{\pi} = 1 - \frac{\text{arctg } x}{\pi}. \end{aligned}$$

У нас зараз виникає певна проблема. Імовірність потрапляння в будь-яку точку нулева. Спробуємо тоді оцінити ймовірність потрапляння в маленький окіл даної точки.

**Definition 2.2.6** Щільністю розподілу неперервної випадкової величини  $\xi$  називають границю

$$f_{\xi}(x) = \lim_{\Delta x \rightarrow 0} \frac{\mathbb{P}\{x \leq \xi < x + \Delta x\}}{\Delta x},$$

якщо така границя існує.

Із означення та властивостей функцій розподілу випливає, що  $f_{\xi}(x) = F'_{\xi}(x)$ .

*Увага! Дане означення не є строгим. Правильніше було би зробити таким чином.*

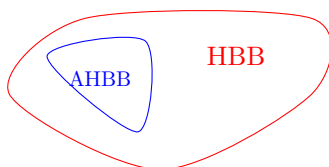
**Definition 2.2.7** Випадкова величина  $\xi$  називається **абсолютно неперервною**, якщо

$$\exists f_{\xi}: \mathbb{R} \rightarrow \mathbb{R} : F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t) dt,$$

водночас функція  $f_{\xi}$  називається **щільністю розподілу**.

**Remark 2.2.8**  $\xi$  – абсолютно неперервна  $\implies \xi$  – неперервна.

На жаль, зворотне твердження не є вірним. Саме тому перше означення про щільність розподілу було нестрогим. Просто тому що не для всіх неперервних випадкових величин можна підібрати щільність розподілу. Це не єдина причина, там воно буде видно нижче.



*TODO: додати приклад*

**Proposition 2.2.9** Властивості щільності розподілу

Справдливе наступне:

- 1)  $f_{\xi} = F'_{\xi}$ ;
- 2)  $f_{\xi} \geq 0$ ;
- 3)  $\int_{-\infty}^{+\infty} f_{\xi}(x) dx = 1$  (умова нормування);
- 4)  $\mathbb{P}\{\xi \in \langle a, b \rangle\} = \int_a^b f_{\xi}(x) dx$ .

**Proof.**

Доведемо кожну властивість:

1) випливає з означення абсолютно неперервної випадкової величини.

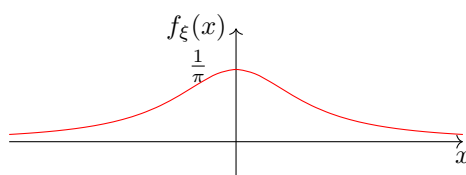
2) Оскільки  $F_\xi$  то звідси випливає  $f_\xi = F'_\xi \geq 0$ .

$$3) \int_{-\infty}^{+\infty} f_\xi(x) dx = \lim_{t \rightarrow +\infty} \int_{-\infty}^t f_\xi(x) dx = \lim_{t \rightarrow +\infty} F_\xi(t) = 1.$$

$$4) \mathbb{P}\{\xi \in \langle a, b \rangle\} = F_\xi(b) - F_\xi(a) = \int_{-\infty}^b f_\xi(x) dx - \int_{-\infty}^a f_\xi(x) dx = \int_a^b f_\xi(x) dx.$$

Всі властивості доведені. ■

**Example 2.2.10** Повернімось до прикладу вище. Ми вже знайшли, чому дорівнює функція розподілу  $F_\xi(x) = 1 - \frac{\operatorname{arcsctg} x}{\pi}$ . Можемо обчислити щільність – отримаємо  $f_\xi(x) = \frac{1}{\pi} \frac{1}{1+x^2}$ . Зараз намалюємо графік та з'ясуємо зміст щільності.

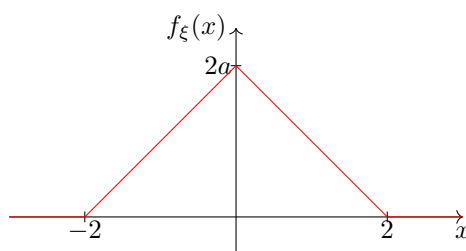


Графік каже наступне. Найчастіше за всього ми будемо потрапляти в окіл точки 0. Це означатиме, з точки зору нашої задачі, що точка перетину з  $OX$  (а це наша  $\xi$ ) буде частіше за всього неподалік від точки 0.

**Example 2.2.11** Нехай випадкова величина  $\xi$  задається щільністю  $f_\xi(x) = \begin{cases} a(2 - |x|), & x \in [-2, 2] \\ 0, & \text{інакше} \end{cases}$ .

Знайти функцію розподілу та ймовірність того, що випадкова величина знаходиться в інтервалі  $(1, 5)$  (перед цією задачею треба з'ясувати, чому дорівнює  $a \in \mathbb{R}$ ).

За умовою нормування,  $\int_{-\infty}^{+\infty} f_\xi(x) dx = 1$ , тобто площа має бути одиничною. Якщо намалювати схематично, то там буде рівнобедрений трикутник, площа якої  $1 = S = \frac{1}{2} \cdot 4 \cdot 2a$ . Отже,  $a = \frac{1}{4}$ .



Знайдемо тепер функцію розподілу, яка рахується як  $F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt$ .

Якщо  $x \in (-\infty, -2]$ , то звідси  $F_\xi(x) = \int_{-\infty}^x 0 dt = 0$ .

Якщо  $x \in (-2, 0]$ , то звідси  $F_\xi(x) = \int_{-\infty}^{-2} f_\xi(t) dt + \int_{-2}^x \frac{1}{4}(2+t) dt = \dots = \frac{x^2}{8} + \frac{x}{2} + \frac{1}{2}$ .

Якщо  $x \in (0, 2]$ , то звідси  $F_\xi(x) = \int_{-\infty}^0 f_\xi(t) dt + \int_0^x \frac{1}{4}(2-t) dt = \left(\frac{0^2}{8} + \frac{0}{2} + \frac{1}{2}\right) + \int_0^x \frac{1}{4}(2-t) dt = \dots = -\frac{x^2}{8} + \frac{x}{2} + \frac{1}{2}$ .

Якщо  $x \in (2, +\infty)$ , то звідси  $F_\xi(x) = \int_{-\infty}^2 f_\xi(t) dt + \int_2^x 0 dt = \left(-\frac{2^2}{8} + \frac{2}{2} + \frac{1}{2}\right) = 1$ .

$$\text{Власне, отримали } F_{\xi}(x) = \begin{cases} 0, & x \in (-\infty, -2] \\ \frac{x^2}{8} + \frac{x}{2} + \frac{1}{2}, & x \in (-2, 0] \\ -\frac{x^2}{8} + \frac{x}{2} + \frac{1}{2}, & x \in (0, 2] \\ 1, & x \in (2, +\infty) \end{cases}.$$

Знайдемо тепер ймовірність. Це можна зробити двома способами, але зроблю це через функцію розподілу, раз її було вже знайдено. Тобто  $\mathbb{P}\{\xi \in (1, 5)\} = F_{\xi}(5) - F_{\xi}(1) = 1 - \frac{7}{8} = \frac{1}{8}$ .

Порада як перевірити, що функція розподілу коректно знайдена:

- 1) продиференціювати назад та отримати ту саму щільність;
- 2) перерірити на неперервність функцію.

## 2.3 Числові характеристики випадкових величин

### 2.3.1 Математичне сподівання

**Definition 2.3.1** Математичним сподіванням (або ще кажуть середнім значенням) випадкової величини  $\xi$  називають таку величину:

$$\mathbb{E}\xi = \sum_{k=1}^{\infty} x_i p_i - \text{для дискретної величини}$$

$$\mathbb{E}\xi = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx - \text{для неперервної величини}$$

Математичне сподівання існує, якщо відбувається абсолютна збіжність (ряда або інтеграла залежно від ситуації).

**Example 2.3.2** Щоб кинути два кубики d6, нам треба заплатити 100 грн. Якщо сума очок складає 7 або 11, то ми виграємо 400 грн. Знайти можливий прибуток за одну гру. Позначимо  $\xi$  – випадкова величина, яка відповідає за суму очок. Запишемо ряд розподілу в даній ситуації.

$\xi$	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Позначимо  $\eta$  – випадкова величина, яка відповідає за прибуток після гри. Ряд розподілу виглядає так:

$\eta$	0	400
$\mathbb{P}$	$\frac{7}{9}$	$\frac{2}{9}$

Зауважимо, що  $\mathbb{P}\{\eta = 400\} = \mathbb{P}\{\xi = 7\} + \mathbb{P}\{\xi = 11\}$ . Порахуємо тепер математичне сподівання:

$$\mathbb{E}\eta = 0 \cdot \frac{7}{9} + 400 \cdot \frac{2}{9} = 88\frac{8}{9}.$$

Тобто це означає, що в будь-якому випадку ми будемо йти в мінус. Платимо 100 грн, а очікується майже 89 грн.

**Example 2.3.3** У нас вже був приклад вище, де знайдена була щільність розподілу  $f_{\xi}(x) = \frac{1}{\pi} \frac{1}{1+x^2}$ .

Тут математичне сподівання не знайти, оскільки  $\mathbb{E}\xi = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx$  – розбіжний інтеграл.

**Definition 2.3.4** Випадкові величини  $\xi_1, \xi_2$ , що задані на одному ймовірнісному просторі, називаються **незалежними**, якщо

$$\forall x, y : \{\xi_1 \leq x\} \text{ та } \{\xi_2 \leq y\} - \text{незалежні події.}$$

Внаслідок чого можемо отримати  $\mathbb{P}\{\xi_1 \leq x, \xi_2 \leq y\} = F_{\xi_1}(x)F_{\xi_2}(y)$ .

**Proposition 2.3.5** Властивості математичних сподівань

Справедливе наступне:

- 1)  $\mathbb{E}c = c$ , під математичним сподіванням мається на увазі стала випадкова величина  $\xi \equiv c$ ;
- 2)  $\mathbb{E}(\lambda\xi) = \lambda\mathbb{E}\xi$ ,  $\lambda \in \mathbb{R}$ ;
- 3)  $\mathbb{E}(\xi + \eta) = \mathbb{E}\xi + \mathbb{E}\eta$ ;
- 4)  $\mathbb{E}(\xi_1\xi_2) = \mathbb{E}\xi_1\mathbb{E}\xi_2$  за умовою, що  $\xi_1, \xi_2$  – незалежні.

**Proof.**

Будемо доводити лише в випадку дискретної випадкової величини.

1) Випадкова величина  $\xi \equiv c$ , що дискретна, означає, що  $\mathbb{P}\{\xi = c\} = 1$ . Внаслідок чого, за означенням математичного сподівання,  $\mathbb{E}c = c \cdot 1 = c$ .

$$2) \mathbb{E}(\lambda\xi) = \sum_{k=1}^{\infty} \lambda x_i p_i = \lambda \sum_{k=1}^{\infty} x_i p_i = \lambda \mathbb{E}\xi.$$

$$3) \mathbb{E}(\xi + \eta) = \sum_{i,j} (x_i + y_j) p_{ij}, \text{ де під ймовірністю } p_{ij} = \mathbb{P}\{\xi + \eta = x_i + y_j\}.$$

$$\text{Звідси випливає, що } \mathbb{E}(\xi + \eta) = x_i \sum_i \sum_j p_{ij} + y_i \sum_i \sum_j p_{ij}.$$

Окремо розглянемо  $\sum_j p_{ij} = \sum_j \mathbb{P}\{\xi + \eta = x_i + y_j\} \stackrel{\text{стверджую}}{=} \mathbb{P}\{\xi = x_i\} = p_i$ . По-перше, зауважимо,

що події  $\{\xi + \eta = x_i + y_j\}, \forall j$  – несумісні уже (це ясно). По-друге, нехай  $\xi(\omega) + \eta(\omega) = x_i + y_{j_0}$ , тоді обов'язково з цього випливає  $\xi(\omega) = x_i$ . Бо якби  $\xi(\omega) \neq x_i$ , то було би  $\eta(\omega) \neq y_j, \forall j$ , що неможливо.

Отже,  $\bigsqcup_j \{\xi + \eta = x_i + y_j\} = \{\xi = x_i\}$ , а далі  $\sigma$ -адитивність ймовірності.

$$\text{Аналогічно доведемо } \sum_i p_{ij} = \sum_i \mathbb{P}\{\xi + \eta = x_i + y_j\} = \mathbb{P}\{\eta = y_j\} = p_j.$$

$$\mathbb{E}(\xi + \eta) = \sum_i x_i p_i + \sum_j y_j p_j = \mathbb{E}\xi + \mathbb{E}\eta.$$

$$\begin{aligned} 4) \mathbb{E}(\xi_1\xi_2) &= \sum_{i,j} (x_i y_j) \mathbb{P}\{\xi_1\xi_2 = x_i y_j\} = \sum_{i,j} (x_i y_j) \mathbb{P}\{\{\xi_1 = x_i\} \cap \{\xi_2 = y_j\}\} \stackrel{\text{незалежність}}{=} \\ &= \sum_{i,j} x_i p_i y_j p_j = \left( \sum_i x_i p_i \right) \left( \sum_j y_j p_j \right) = \mathbb{E}\xi_1 \mathbb{E}\xi_2. \end{aligned}$$

Для неперервних випадкових величин доведення 3), 4) буде пізніше. ■

**2.3.2 Дисперсія, середньоквадратичне відхилення**

Задана випадкова величина  $\xi$ . Суть дисперсії полягає в тому, щоб охарактеризувати степінь розкиду значень  $\xi$  навколо середнього  $\mathbb{E}\xi$ .

**Definition 2.3.6** Дисперсією випадкової величини  $\xi$  називають таку величину:

$$\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2$$

**Remark 2.3.7** Випадкова величина  $\xi - \mathbb{E}\xi$  – це розкидані значення навколо нуля. Якщо будемо брати просто  $\mathbb{E}(\xi - \mathbb{E}\xi)$ , то це буде завжди нуль (через властивості можна довести), тому що є додатні значення та від'ємні, та усереднене буде нулевим. Тому ми додали другий степінь, щоб можна було з обох сторін брати додатні числа.

Розпишемо дисперсію більш детально:

$$\mathbb{D}\xi = \mathbb{E}(\xi^2) - 2\xi\mathbb{E}\xi + (\mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - \mathbb{E}(2\xi\mathbb{E}\xi) + \mathbb{E}(\mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - 2(\mathbb{E}\xi)^2 + (\mathbb{E}\xi)^2.$$

Отже, дисперсія випадкової величини  $\xi$  можна обчислити за такою формулою:

$$\mathbb{D}\xi = \mathbb{E}(\xi^2) - (\mathbb{E}\xi)^2$$

**Proposition 2.3.8** Властивості дисперсії

Справедливе наступне:

- 1)  $\mathbb{D}\xi \geq 0$ ;
- 2)  $\mathbb{D}\xi = 0 \iff \xi \equiv c$ ;
- 3)  $\mathbb{D}(c\xi) = c^2\mathbb{D}\xi$ ;
- 4)  $\mathbb{D}(\xi_1 + \xi_2) = \mathbb{D}\xi_1 + \mathbb{D}\xi_2$  за умовою, що  $\xi_1, \xi_2$  – незалежні.
- 1) впливає з означення математичного сподівання та решта 2), 3), 4), 5) впливають з другої формули.

**Definition 2.3.9** Середньоквадратичним відхиленням випадкової величини  $\xi$  називають наступне:

$$\sigma_\xi = \sqrt{\mathbb{D}\xi}$$

Це ще часто називають **стандартним відхиленням**. Характеризує, по суті, наскільки від середнього значення ми можемо відхилитися ліворуч чи праворуч.

**Example 2.3.10** Позначимо  $\xi$  – випадкова величина, яка описує кількість забитих м'ячів командою МЮ за останні 33 матчі АПЛ сезону 23/24. На основі відомих даних знайти математичне сподівання та дисперсію.

У мене вийшов ось такий ряд розподілу, поки скролив усі матчі:

$\xi$	0	1	2	3	4
$\mathbb{P}$	$\frac{7}{33}$	$\frac{10}{33}$	$\frac{9}{33}$	$\frac{5}{33}$	$\frac{2}{33}$

Обчислимо дисперсію випадкової величини. Для цього спочатку знайдемо математичне сподівання та ще додатковий параметр.

$$\mathbb{E}\xi = 0 \cdot \frac{7}{33} + 1 \cdot \frac{10}{33} + 2 \cdot \frac{9}{33} + 3 \cdot \frac{5}{33} + 4 \cdot \frac{2}{33} = \frac{51}{33} = 1\frac{18}{33} \text{ (ну коротше, десь півтори м'ячи в середньому).}$$

$$\mathbb{E}\xi^2 = 0^2 \cdot \frac{7}{33} + 1^2 \cdot \frac{10}{33} + 2^2 \cdot \frac{9}{33} + 3^2 \cdot \frac{5}{33} + 4^2 \cdot \frac{2}{33} = \frac{123}{33} = 3\frac{24}{33}.$$

Отже, звідси  $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 \approx 1.33$ . Відхилення  $\sigma_\xi = \sqrt{\mathbb{D}\xi} \approx 1.15$ . Тобто очікується зазвичай десь від 0.38 до 2.7 голів.

## 2.4 Коротко про генератрису в теорії ймовірностей

Припустимо, що випадкова величина  $\xi \in \{0, 1, 2, \dots\}$ , на кожному значенні ймовірність відповідно  $p_0, p_1, p_2, \dots$ . Ми вже знаємо, що генератриса (generating function) послідовності  $\{a_n, n \geq 0\}$  визначається як  $G(a_n; z) = \sum_{n=0}^{\infty} a_n z^n$ . Тепер візьмемо послідовність ймовірностей  $\{p_n, n \geq 0\}$  – отримаємо

$$\text{генератрису } G(p_n; z) = \sum_{n=0}^{\infty} p_n z^n \stackrel{\text{насправді}}{=} \mathbb{E}(z^\xi).$$

Отже, маємо таке означення:

**Definition 2.4.1** Генератрисою випадкової величини  $\xi \in \{0, 1, 2, \dots\}$  називають наступне:

$$G_\xi(z) = \mathbb{E}(z^\xi)$$

Англійською це називають **probability generating function**.

**Proposition 2.4.2** Властивості генератрису випадкової величини

Справедливе наступне:

- 1)  $G_\xi(z)$  визначений при  $|z| \leq 1$  (ряд при таких  $z$  рівномірно збіжний);
- 2)  $G_\xi(1) = 1$ ;
- 3)  $G_\xi$  – неспадна;
- 4)  $G_\xi \geq 0$ ;
- 5) Розподіл випадкової величини  $\xi$  повністю визначається генератрисою  $G_\xi(t)$ ;
- 6) Нехай  $\xi, \eta$  – незалежні. Тоді  $G_{\xi+\eta}(z) = G_\xi(z) \cdot G_\eta(z)$ ;

**Proof.**

Доведемо окремо:

1)  $|z^k p_k| \leq p_k$  при наших обраних  $z$ , а тому якщо розглянути ряд  $\sum_{k=0}^{\infty} p_k = 1$  (який збіжний), то за

мажорантною ознакою Ваєрштраса,  $G_{\xi}(z) = \sum_{k=0}^{\infty} z_k p^k$  буде рівномірно збіжний при  $|z| \leq 1$ .

2)  $G_{\xi}(1) = \sum_{k=0}^{\infty} p_k = 1$ .

3) Зрозуміло.

4) Зрозуміло.

5) Дана властивість означає, що за допомогою генератрис ми можемо повністю відновити ряд розподілу. Тобто кожному значенню випадкової величини можемо відновити ймовірність. Доведемо, що  $p_k = \frac{G_{\xi}^{(k)}(0)}{k!}$ . Для цього ми просто диференціюємо (скільки треба), маємо права в силу того, що це степеневий ряд. А далі просто підставляємо  $z = 0$ .

6)  $G_{\xi+\eta}(z) = \mathbb{E}(z^{\xi+\eta}) = \mathbb{E}(z^{\xi} z^{\eta}) = \mathbb{E}(z^{\xi}) \mathbb{E}(z^{\eta}) = G_{\xi}(z) G_{\eta}(z)$ .

Єдине треба переконатися, що коли  $\xi, \eta$  незалежні, то звідси  $z^{\xi}, z^{\eta}$  теж незалежні. Це вже поясню, коли буду вводити строго через теорію міри.

Всі властивості (майже чесно) довели. ■

**Remark 2.4.3** Зауважимо, що  $\mathbb{E}\xi = G'_{\xi}(1)$  та  $\mathbb{E}\xi^2 = G'_{\xi}(1) + G''_{\xi}(1)$ . Це дозволить простіше обчислювати дисперсію.

## 2.5 Популярні розподіли дискретних випадкових величин

### 2.5.1 Розподіл Бернуллі

Розподіл Бернуллі можна описати таким рядом розподілу:

$\xi$	0	1
$\mathbb{P}$	$q$	$p$

У цьому випадку  $p + q = 1$ .

Тобто розподіл Бернуллі характеризує успіх виконання даної події чи невдачі.

Функція розподілу даної випадкової величини буде виглядати таким чином:

$$F_{\xi}(x) = \begin{cases} 0, & x < 0 \\ q, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

Якщо обчислити математичне сподівання та дисперсію, то отримаємо таку формулу:

$$\mathbb{E}\xi = p, \quad \mathbb{D}\xi = pq.$$

### 2.5.2 Біноміальний розподіл

Біноміальний розподіл можна описати таким рядом розподілу:

$\xi$	0	1	...	$k$	...	$n$
$\mathbb{P}$	$q^n$	$npq^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$p^n$

Біноміальний розподіл описує кількість успіхів у схемі Бернуллі з  $n$  випробувань та з імовірністю успіху  $p$ . Такий ряд розподілу можна ось так коротко записати:  $\xi \sim \text{Bin}(n, p)$ .

Мабуть, писати функцію розподілу такої випадкової величини не буду, бо тут ясно цілком, що буде.

Краще обчислю детально математичне сподівання разом із дисперсією.

Оскільки  $\xi \in \{0, 1, 2, \dots\}$ , то ми можемо все це зробити через генератрис.

$$G_{\xi}(z) = \sum_{k=0}^n p_k z^k = \sum_{k=0}^n C_n^k p^k q^{n-k} z^k = \sum_{k=0}^n C_n^k (pz)^k q^{n-k} = (pz + q)^n.$$

Зокрема звідси  $G'_\xi(z) = pn(pz+q)^{n-1}$ , а також  $G''_\xi(z) = p^2n(n-1)(pz+q)^{n-2}$ . Тепер можемо спокійно знайти бажані параметри.

$$\mathbb{E}\xi = pn(p \cdot 1 + q) = pn.$$

$$\mathbb{D}\xi = pn + p^2n(n-1) - p^2n^2 = np + n^2p^2 - n^2p^2 - np^2 = np(1-p) = npq.$$

Отже, отримали:

$$\mathbb{E}\xi = np, \quad \mathbb{D}\xi = npq.$$

### 2.5.3 Геометричний розподіл

Геометричний розподіл можна описати таким рядом розподілу:

$\xi$	0	1	...	k	...
$\mathbb{P}$	p	qp	...	$q^k p$	...

Геометричний розподіл описує кількість невдач до першого успіху в схемі Бернуллі з ймовірністю успіху  $p$ . Такий розподіл можна ось так коротко записати:  $\xi \sim \text{Geom}_0(p)$ .

Як задати функцію розподілу, тут теж не виникає питань. Тому порахуємо параметри.

Знову оскільки  $\xi \in \{0, 1, \dots\}$ , ми можемо застосувати генератрису для обчислення.

$$G_\xi(z) = \sum_{k=0}^n p_k z^k = \sum_{k=0}^{\infty} q^k p z^k = p \sum_{k=0}^{\infty} (qz)^k = \frac{p}{1-qz}.$$

Зауважимо, що ми останню рівність можемо допустити, якщо  $|qz| < 1$ . Але така нерівність завжди виконується. У нас  $|z| \leq 1$ , при цьому  $q \in (0, 1)$ , тому звідси  $|qz| \leq q < 1$ .

Зокрема  $G'_\xi(z) = pq(1-pz)^{-2}$ , а також  $G''_\xi(z) = 2pq^2(1-pz)^{-3}$ . Тепер знайдемо бажані параметри.

$$\mathbb{E}\xi = \frac{q}{1-q} = \frac{q}{p};$$

$$\mathbb{D}\xi = \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2} = \frac{qp + 2q^2 - q^2}{p^2} = \frac{qp + q^2}{p^2} = \frac{q}{p^2}.$$

Отже, отримали:

$$\mathbb{E}\xi = \frac{q}{p}, \quad \mathbb{D}\xi = \frac{q}{p^2}.$$

**Example 2.5.1** Підкидаємо монету, доки не випаде решка. З'ясувати, коли ми зазвичай закінчимо кидати.

Маємо  $\xi \sim \text{Geom}_0\left(\frac{1}{2}\right)$ , де ймовірність  $\frac{1}{2}$  означає ймовірність успіху (випадіння решки). Тоді за

щойно отриманими результатами,  $\mathbb{E}\xi = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1$ . Тобто ми можемо або миттєво закінчити, або через одне підкидання.

### 2.5.4 Розподіл Пуассона

Розподіл Пуассона можна описати таким рядом розподілу:

$\xi$	0	1	2	...	k	...
$\mathbb{P}$	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2}$	...	$\frac{\lambda^k e^{-\lambda}}{k!}$	...

У цьому випадку параметр  $\lambda > 0$ . Такий розподіл можна ось так коротко записати:  $\xi \sim \text{Pois}(\lambda)$ .

Знову ж таки, можемо через генератрису знайти наші улюблені параметри.

$$G_\xi(z) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} z^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}.$$

Зокрема якщо порахувати першу та другу похідні в одиниці, то отримаємо:

$$\mathbb{E}\xi = \lambda, \quad \mathbb{D}\xi = \lambda.$$

Щоб поговорити про застосування, треба трошки затронутися такою річ. Є таке поняття як **потік** – це послідовність подій, що настають в певні моменти часу одна за одною. Потік характеризується випадковою величиною  $\xi(t)$  – кількістю подій, що наступили протягом проміжку часу  $[0, t)$ .

Потоків бувають декілька, які розглядаються переважно в інших дисциплінах математики. Ми зосередимося лише на однойменному потоку.



## Потік Пуассона

**Definition 2.5.2** Потоком Пуассона називають потік із такими обмеженнями:

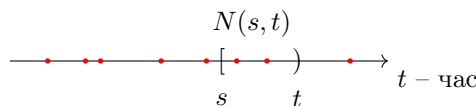
1. **Стационарність.** Кількість подій, що настають за певний проміжок часу, залежить лише від довжини проміжку та не залежить від того, де цей проміжок розташований на часі.
2. **Відсутність післядії.** Якщо проміжки часу неперетинні, то кількість подій, які за ці проміжки відбулися, є незалежними подіями.
3. **Ординарність.** Події настають поодиночі.

**Example 2.5.3** Життєвий приклад нестационарного потоку. Уявімо собі кількість людей, які проходять через турнікети в київському метро. Наприклад, візьмемо одногодинний проміжок. Умовно о 7-8 годинах кількість людей (подія – прохід людини через турнікет) буде явно більшою, ніж умовно о 12-13.

**Example 2.5.4** Життєвий приклад потоку, де не працює відсутність післядії. Уявімо собі заклад "Тайський привіт", куди ходять приблизно однакова кількість в будь-який проміжок часу. Потім заходить Клопотенко та починає розповідати, який заклад неперевершений. Після цього кількість подій (прийти поїсти там) суттєво збільшиться в силу хайпа.

**Example 2.5.5** Життєвий приклад неординарного потоку. Знову ж таки, київське метро, нема охорони. Одна людина проходить через турнікет чесним шляхом, а тут миттєво інша людина проходить безкоштовно через цей турнікет (може перестрибнути або пролізти, як завгодно).

Тепер зараз я трохи математизую те, що написав. Позначимо  $N(s, t)$  – випадкова величина, яка характеризує кількість подій на проміжку  $[s, t)$ .



Математично опишемо кожне обмеження, яке ми наклали на потік Пуассона:

1. **Стационарність.**  $N(s, t) \stackrel{d}{=} N(s+h, t+h)$  при  $h > 0$ . Рівність  $\stackrel{d}{=}$  означає рівність за розподілом. Це означає іншими словами, що  $\mathbb{P}\{N(s, t) = k\} = \mathbb{P}\{N(s+h, t+h) = k\}, \forall k \geq 0$ .
2. **Відсутність післядії.**  $N(s_1, t_1), N(s_2, t_2), \dots, N(s_n, t_n)$  – незалежні випадкові величини в сукупності, тут  $s_1 < t_1 < s_2 < t_2 < \dots < s_n < t_n$ .
3. **Ординарність.**  $\mathbb{P}\{N(t, t+h) \geq 2\} = o(h)$  при  $h \rightarrow 0$ .

**Theorem 2.5.6** Для будь-якого непорожнього потоку Пуассона існує число  $\lambda > 0$  – так звана інтенсивність потоку, – для якого  $\forall s \leq t: N(s, t) \sim \text{Pois}(\lambda(t-s))$ .

Мабуть, я дану теорему розіб'ю на кілька лем, щоб можна було легше сприймати.

**Lemma 2.5.7** Для будь-якого непорожнього потоку Пуассона  $\mathbb{P}\{N(0, t) = 0\} = p^t$  для всіх  $t > 0$ . Ми тут позначили за  $p = \mathbb{P}\{N(0, 1) = 0\}$ , причому слід зазначити, що  $p \in (0, 1)$ .

**Proof.**

Для початку нам буде потрібна ось така рівність:

$$\begin{aligned} \mathbb{P}\{N(0, nt) = 0\} &= \mathbb{P}\{N(0, t) = 0, N(t, 2t) = 0, \dots, N((n-1)t, nt) = 0\} \stackrel{\text{відсутність післядії}}{=} \\ &= \mathbb{P}\{N(0, t) = 0\} \mathbb{P}\{N(t, 2t) = 0\} \dots \mathbb{P}\{N((n-1)t, nt) = 0\} \stackrel{\text{стационарність}}{=} \mathbb{P}^n\{N(0, t) = 0\}. \end{aligned}$$

Позначимо  $\mathbb{P}\{N(0, 1) = 0\} = p$ . Тоді за щойно доведеним,  $\mathbb{P}\{N(0, n) = 0\} = p^n, n \in \mathbb{N}$ .

Розглянемо тепер ймовірність того, що кількість подій на  $\left[0, \frac{1}{n}\right)$  нулева. Тоді

$$\mathbb{P}\left\{N\left(0, \frac{1}{n}\right) = 0\right\} \stackrel{\text{за вище доведеним}}{=} \mathbb{P}^{\frac{1}{n}}\{N(0, 1) = 0\} = p^{\frac{1}{n}}.$$

$$\text{Зокрема отримаємо } \mathbb{P}\left\{N\left(0, \frac{m}{n}\right) = 0\right\} = \mathbb{P}^m\left\{N\left(0, \frac{1}{n}\right) = 0\right\} = p^{\frac{m}{n}}.$$

Висновок:  $\mathbb{P}\{N(0, r) = 0\} = p^r$  для кожного додатного  $r \in \mathbb{Q}$ .

Але ми хочемо для будь-якого проміжку часу формату  $[0, t)$ , тобто  $t \in \mathbb{R}$  додатне.

Розглянемо  $t \in \mathbb{R}_{>0}$ . Для неї можна підібрати дві монотонні послідовності додатних раціональних чисел: одна послідовність  $\{r_i^+, i \geq 1\}$  праворуч спадає до  $t$ , друга послідовність  $\{r_i^-, i \geq 1\}$  ліворуч зростає до  $t$ . Зауважимо, що справедлива нерівність:

$$\mathbb{P}\{N(0, r_i^+) = 0\} \leq \mathbb{P}\{N(0, t) = 0\} \leq \mathbb{P}\{N(0, r_i^-) = 0\}.$$

$$p^{r_i^+} \leq \mathbb{P}\{N(0, t) = 0\} \leq p^{r_i^-}.$$

Якщо спрямувати  $i \rightarrow \infty$  та зауважити, що  $\mathbb{P}\{N(0, t) = 0\}$  – стаціонарна штука (тобто не змінюється при зміні  $i$ ), то, урахувавши той факт, що  $r_i^+, r_i^- \rightarrow t$ , отримаємо  $p^{r_i^+}, p^{r_i^-} \rightarrow t$ , внаслідок чого  $\mathbb{P}\{N(0, t) = 0\} = p^t$ .

Висновок:  $\mathbb{P}\{N(0, t) = 0\} = p^t$  для кожного  $t > 0$ .

Треба ще зазначити, що ймовірність  $p \in (0, 1)$ , зараз пояснимо недопустимість крайніх значень.

При  $p = 1$  потік Пуассона стане порожнім, а це суперечить початковим умовам.

При  $p = 0$  ситуація набагато цікавіша. Зауважимо, що  $\mathbb{P}\{N(s, t) = 0\} \stackrel{\text{стаціонарність}}{=} \mathbb{P}\{N(0, t-s) = 0\} = 0^{t-s} = 0$ , але тоді звідси  $\mathbb{P}\{N(s, t) \geq 2\} = 1$ . Тобто порушується умова ординарності – не те, що нам треба. ■

**Lemma 2.5.8** Для будь-якого непорожнього потоку Пуассона існує  $\lambda > 0$ , для якого  $\mathbb{P}\{N(s, t) = 0\} = e^{-\lambda(t-s)}$ .

**Proof.**

Дійсно, раз  $p \in (0, 1)$ , то ми можемо покласти  $\lambda = -\ln p$ . Тоді звідси

$$\mathbb{P}\{N(s, t) = 0\} \stackrel{\text{стаціонарність}}{=} \mathbb{P}\{N(0, t-s) = 0\} = p^{t-s} = e^{-\lambda(t-s)}.$$

■

**Lemma 2.5.9**  $N(s, t) \sim \text{Pois}(\lambda(t-s))$ , тобто  $\mathbb{P}\{N(s, t) = k\} = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}$ .

**Proof.**

Щойно при  $k = 0$  ми цю рівність довели вище. Тому надалі беремо  $k \geq 1$ . Ми почнемо з найпростішого та доведемо випадок  $\mathbb{P}\{N(0, t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ . Для спрощення позначу за  $\mathbb{P}\{N(0, t) = s\} = p_s(t)$ , де  $s \geq 0$ .

Ми вже маємо, що  $p_0(t) = e^{-\lambda t} \stackrel{\text{Тейлор}}{=} 1 - \lambda t + o(t)$  при  $t \rightarrow 0 + 0$ .

Таким чином,  $p_1(t) = 1 - \mathbb{P}\{N(0, t) = 0\} - \mathbb{P}\{N(0, t) \geq 2\} \stackrel{\text{ординарність}}{=} \lambda t + o(t)$  при  $t \rightarrow 0 + 0$ .

Розглянемо тепер  $p_k(t+h)$  та з'ясуємо, що це буде.

$$p_k(t+h) = \mathbb{P}\{N(0, t+h) = k\} = \mathbb{P}\{N(0, t) = k, N(t, t+h) = 0\} + \mathbb{P}\{N(0, t) = k-1, N(t, t+h) = 1\} + \mathbb{P}\{N(0, t) = k-l, N(t, t+h) = l, l \geq 2\} \stackrel{\text{відсутність післядії} + \text{стаціонарність}}{=}$$

$$= \mathbb{P}\{N(0, t) = k\} \mathbb{P}\{N(0, t) = 0\} + \mathbb{P}\{N(0, t) = k-1\} \mathbb{P}\{N(0, t) = 1\} + \sum_{l=2}^k \mathbb{P}\{N(0, t) = k-l\} \mathbb{P}\{N(0, t) = l\}$$

$$\stackrel{\text{ординарність}}{=} p_k(t)p_0(h) + p_{k-1}(t)p_1(h) + o(h).$$

$$p_k(t+h) = (1 - \lambda h + o(h))p_k(t) + (\lambda h + o(h))p_{k-1}(t) + o(h).$$

$$p_k(t+h) - p_k(t) = -\lambda h p_k(t) + \lambda h p_{k-1}(t) + o(h).$$

$$\frac{p_k(t+h) - p_k(t)}{h} = -\lambda p_k(t) + \lambda p_{k-1}(t) + \frac{o(h)}{h}.$$

Нарешті, спрямуємо  $h \rightarrow 0 + 0$  – врешті решт отримаємо нескінченну систему диференціальних рівнянь з деякими умовами:

$$\begin{cases} p'_k(t) = -\lambda p_k(t) + \lambda p_{k-1}(t), & k \geq 1 \\ p_0(t) = e^{-\lambda t} \\ p_k(0) = 0 \end{cases}.$$

Будемо шукати розв'язок такого вигляду:  $p_k(t) = e^{-\lambda t} v_k(t)$ . Похідна  $p'_k(t) = e^{-\lambda t} v'_k(t) - \lambda e^{-\lambda t} v_k(t)$ .

Всі ці данні підставимо в систему диференціальних рівнянь – отримаємо простий вигляд:

$$\begin{cases} v'_k(t) = \lambda v_{k-1}(t), & k \geq 1 \\ v_0(t) = 1 \\ v_k(0) = 0 \end{cases}.$$

Далі, граючись із рекурентним співвідношенням, отримаємо  $v_k(t) = \frac{(\lambda t)^k}{k!}$ . У результаті чого, ми

довели, що  $p_k(t) = \mathbb{P}\{N(0, t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ . Нарешті,

$$\mathbb{P}\{N(s, t) = k\} \stackrel{\text{стаціонарність}}{=} \mathbb{P}\{N(0, t-s) = k\} = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}.$$

**Example 2.5.10** На станцію швидкої допомоги за годину в середньому доходить 90 викликів. З'ясувати, яка ймовірність того, що за 1 хвилину буде принаймні 2 виклики.

Зауважимо, що це описує потік Пуассона, тому  $N(s, t) \sim \text{Pois}$ . Якщо за 1 годину в нас 90 викликів в середньому, то за 1 хвилину в нас 1.5 викликів в середньому (за стаціонарністю).

Отже, введемо випадкову величину  $\xi$  – кількість викликів протягом 1 хвилини. Тоді  $\xi \in \text{Pois}(1.5)$ , оскільки  $\mathbb{E}\xi = 1.5 = \lambda$ . Внаслідок чого, отримаємо:

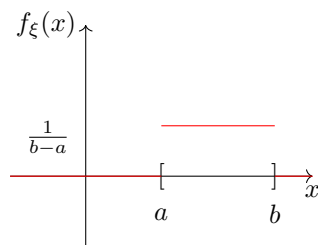
$$\mathbb{P}\{\xi \geq 2\} = 1 - \mathbb{P}\{\xi = 0\} - \mathbb{P}\{\xi = 1\} = 1 - 2.5e^{-1.5}.$$

## 2.6 Популярні розподіли абсолютно неперервних випадкових величин

### 2.6.1 Рівномірний розподіл

Рівномірний розподіл можна описати такою щільністю розподілу:

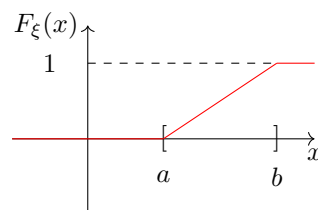
$$f_{\xi}(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & \text{інакше} \end{cases}$$



Такий розподіл можна ось так коротко записати:  $\xi \sim U(a, b)$  при  $a, b \in \mathbb{R}$ , для яких  $a < b$ .

Можна спробувати самостійно обчислити функцію розподілу, математичне сподівання та дисперсію – отримаємо наступне:

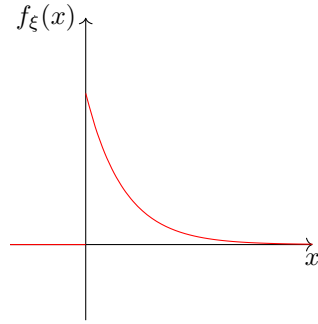
$$F_{\xi}(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & x > b \end{cases}, \quad \mathbb{E}\xi = \frac{a+b}{2}, \quad \mathbb{D}\xi = \frac{(b-a)^2}{12}.$$



### 2.6.2 Експоненційний розподіл

Експоненційний розподіл можна описати такою щільністю розподілу:

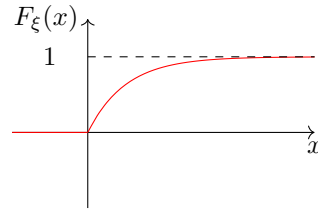
$$f_{\xi}(X) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{інакше} \end{cases}$$



Такий розподіл можна ось так коротко записати:  $\xi \in \text{Exp}(\lambda)$  при  $\lambda > 0$ .

Можна спробувати самостійно обчислити функцію розподілу, математичне сподівання та дисперсію – отримаємо наступне:

$$F_{\xi}(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\lambda x}, & x > 0 \end{cases}, \quad \mathbb{E}\xi = \frac{1}{\lambda}, \quad \mathbb{D}\xi = \frac{1}{\lambda^2}.$$

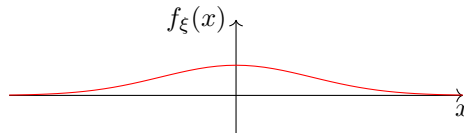


### 2.6.3 Нормальний розподіл

I. *Стандартний гаусівський розподіл.*

Розподіл Гауса визначається таким чином:

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$



Такий розподіл можна ось так коротко записати:  $\xi \sim \mathcal{N}(0, 1^2)$ . Ці числа всередині дужок мають певний сенс, про який згодом обговоримо.

Зауважимо, що  $\mathbb{E}\xi = 0$ . Якщо коротко, то тут вступає непарність підінтегральної функції. Зосередимося на дисперсії

$$\begin{aligned} \mathbb{D}\xi &= \mathbb{E}\xi^2 = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx = 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx \stackrel{\frac{x^2}{2}=t}{=} \sqrt{\frac{2}{\pi}} \int_0^{+\infty} 2te^{-t} \frac{dt}{\sqrt{2t}} = \\ &= \frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{\frac{1}{2}} e^{-t} dt = \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) = 1. \end{aligned}$$

Отже, наспраді кажучи,  $0, 1^2$  – це відповідно математичне сподівання та дисперсія,  $\xi \in \mathcal{N}(\mathbb{E}\xi, \sigma_{\xi}^2)$ . Щодо функції розподілу тут доволі важка ситуація. Тим не менш, спочатку розглянемо  $x \geq 0$ .

$$F_{\xi}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{2} + \Phi(x).$$

**Definition 2.6.1** Функцією Лапласа назовемо таку функцію при  $x \in \mathbb{R}$ :

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

### Proposition 2.6.2 Властивості функції Лапласа

Справедливе наступне:

1)  $\Phi$  – непарна функція;

2)  $\lim_{x \rightarrow \pm\infty} \Phi(x) = \pm \frac{1}{2}$ .

*Вправа: довести.*

Повертаючись до теми функції розподілу, ми отримали  $F_\xi(x) = \frac{1}{2} + \Phi(x)$  при  $x \geq 0$ . Маючи властивості, отримаємо  $F_\xi(-x) = \frac{1}{2} - \Phi(x)$ .

Висновок: для випадкової величини  $\xi \sim \mathcal{N}(0, 1^2)$  маємо наступне:

$$F_\xi(x) = \frac{1}{2} + \Phi(x), \quad \mathbb{E}\xi = 0, \quad \mathbb{D}\xi = 1^2.$$

II. *Загальний гаусівський розподіл.*

Випадкову величину  $\xi$  задамо ось таким чином:

$$\xi = a + \sigma\xi_0, \quad \xi_0 \sim \mathcal{N}(0, 1^2)$$

Такий розподіл можна ось так коротко записати:  $\xi \sim \mathcal{N}(a, \sigma^2)$ .

$$\mathbb{E}\xi = \mathbb{E}(a + \sigma\xi_0) = a + \sigma\mathbb{E}\xi_0 = a;$$

$$\mathbb{D}\xi = \mathbb{D}(a + \sigma\xi_0) = \sigma^2\mathbb{D}\xi_0 = \sigma^2.$$

## 3 Випадкові вектори

### 3.1 Перші означення

**Definition 3.1.1** Задано  $(\Omega, \mathcal{F})$  – вимірний простір, також розглянемо  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

Функція  $\vec{\xi}: \Omega \rightarrow \mathbb{R}^n$  називається **випадковим вектором**, якщо

$$\xi_i, i = \overline{1, n} - \text{випадкова величина.}$$

**Definition 3.1.2 (Сумісною) функцією розподілу** випадкового вектора  $\vec{\xi}$  називають таку функцію:

$$F_{\vec{\xi}}(\vec{x}) = \mathbb{P}\{\xi_1 \leq x_1, \dots, \xi_n \leq x_n\}, \vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$$

Часто ми будемо розглядати випадкові вектори в двовимірному (інколи тривірному) просторах.

#### Proposition 3.1.3 Властивості функції розподілу

Справедливе наступне:

- 1)  $F_{\vec{\xi}} \in [0, 1]$ ;
- 2)  $F_{\vec{\xi}}$  неспадна функція за кожним аргументом;
- 3)  $F_{\vec{\xi}}$  неперервна праворуч за кожним аргументом;
- 4)  $\lim_{\text{хоча б один із } x_i \rightarrow -\infty} F_{\vec{\xi}}(x_1, \dots, x_i, \dots, x_n) = 0, \quad \lim_{\text{всі } x_i \rightarrow +\infty} F_{\vec{\xi}}(x_1, \dots, x_i, \dots, x_n) = 1$ ;
- 5)  $\lim_{\text{всі } x_i \rightarrow +\infty, \text{ крім } x_k} F_{\vec{\xi}}(x_1, \dots, x_k, \dots, x_n) = F_{\xi_k}(x_k)$ ;
- 6)  $\forall \{i_1, \dots, i_k\} \subset \{1, \dots, n\} : \lim_{x_{i_t} \rightarrow +\infty, t=\overline{1, k}} F_{\vec{\xi}}(\vec{x}) = F_{\xi_{i_1} \dots \xi_{i_k}}(x_{i_1}, \dots, x_{i_k})$ ;
- 7) Нехай координати  $\xi_1, \dots, \xi_n$  незалежні в сукупності. Тоді  $F_{\vec{\xi}}(\vec{x}) = F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n)$ ;

Окремо властивість для двовимірного вектора.

8) Позначимо прямокутник  $\Pi = (a, b] \times (c, d]$ . Тоді  $\mathbb{P}\{\vec{\xi} \in \Pi\} = F_{\vec{\xi}}(b, d) - F_{\vec{\xi}}(b, c) - F_{\vec{\xi}}(a, d) + F_{\vec{\xi}}(a, c)$ .  
Всі властивості доводяться аналогічно, як це було при функціях розподілу випадкової величини.  
Хіба що обговорю останню властивість.

#### Proof.

Зауважимо, що виконується така рівність подій:

$$\{\xi_1 \leq b, \xi_2 \leq d\} = \{\xi_1 \leq a, \xi_2 \leq c\} \sqcup \{a < \xi_1 \leq b, \xi_2 \leq c\} \sqcup \{\xi_1 \leq a, c < \xi_2 \leq d\} \sqcup \{\vec{\xi} \in \Pi\}.$$

Переходячи до ймовірностей, отримаємо таке:

$$F_{\vec{\xi}}(b, d) = F_{\vec{\xi}}(a, c) + \mathbb{P}\{a < \xi_1 \leq b, \xi_2 \leq c\} + \mathbb{P}\{\xi_1 \leq a, c < \xi_2 \leq d\} + \mathbb{P}\{\vec{\xi} \in \Pi\}.$$

Зауважимо, що  $\mathbb{P}\{a < \xi_1 \leq b, \xi_2 \leq c\} = F_{\vec{\xi}}(b, c) - F_{\vec{\xi}}(a, c)$ .

Зауважимо, що  $\mathbb{P}\{\xi_1 \leq a, c < \xi_2 \leq d\} = F_{\vec{\xi}}(a, d) - F_{\vec{\xi}}(a, c)$ .

Зокрема звідси  $\mathbb{P}\{\vec{\xi} \in \Pi\} = F_{\vec{\xi}}(b, d) - F_{\vec{\xi}}(b, c) - F_{\vec{\xi}}(a, d) + F_{\vec{\xi}}(a, c)$ . ■

**Definition 3.1.4** Функції розподілу  $F_{\xi_k}(x_k)$ , які були отримані із сумісної функції розподілу  $F_{\vec{\xi}}(\vec{x})$  за властивістю 5), називаються **маргинальними функціями розподілу**.

### 3.2 Основні види випадкових векторів

#### 3.2.1 Дискретний випадковий вектор

**Definition 3.2.1** Випадковий вектор  $\vec{\xi}$  називається **дискретним**, якщо

$$\xi_i, i = \overline{1, n} - \text{дискретна випадкова величина.}$$

Розглянемо поки двовимірний дискретний випадковий вектор  $\vec{\xi} = (\xi_1, \xi_2)^T$ . Для того, щоб задати дискретний випадковий вектор, крім значень векторів ще треба мати ймовірності, із якими ці значення приймаються.

$$p_{ij} = \mathbb{P}\{\xi_1 = x_i, \xi_2 = y_j\}, \quad \sum_{i,j} p_{ij} = 1.$$

Часто всі дані зручно записувати в **таблицю розподілу**, що виглядає таким чином:

$\xi_1/\xi_2$	$y_1$	$y_2$	$\dots$	$y_j$	$\dots$
$x_1$	$p_{11}$	$p_{12}$	$\dots$	$p_{1j}$	$\dots$
$x_2$	$p_{21}$	$p_{22}$	$\dots$	$p_{2j}$	$\dots$
$\vdots$	$\dots$	$\dots$	$\ddots$	$\dots$	$\ddots$
$x_i$	$p_{i1}$	$p_{i2}$	$\dots$	$p_{ij}$	$\dots$
$\vdots$	$\dots$	$\dots$	$\ddots$	$\dots$	$\ddots$

Зазвичай обмежують скінченними таблицями, але тим не менш.

Зауважимо ось таку цікаву особливість:

$$\mathbb{P}\{\xi_1 = x_i\} = \mathbb{P}\left\{\bigcup_{j=1}^{\infty}\{\xi_1 = x_i, \xi_2 = y_j\}\right\} = \sum_{j=1}^{\infty} \mathbb{P}\{\xi_1 = x_i, \xi_2 = y_j\} = \sum_{j=1}^{\infty} p_{ij}.$$

$$\mathbb{P}\{\xi_2 = y_j\} \stackrel{\text{аналогічно}}{=} \sum_{i=1}^{\infty} p_{ij}.$$

Тобто щоб знайти ймовірність  $\mathbb{P}\{\xi_1 = x_i\} = p_i$ , треба просумувати  $j$ -ий стовпчик таблиці. Отримаємо так званий **маргинальний ряд розподілу**:

$\xi_1$	$x_1$	$x_2$	$\dots$	$x_i$	$\dots$
$\mathbb{P}$	$p_1$	$p_2$	$\dots$	$p_i$	$\dots$

Аналогічно щоб знайти ймовірність  $\mathbb{P}\{\xi_2 = y_j\} = q_j$ , треба просумувати  $i$ -ий рядок таблиці. Отримаємо теж так званий **маргинальний ряд розподілу**:

$\xi_2$	$y_1$	$y_2$	$\dots$	$y_j$	$\dots$
$\mathbb{P}$	$q_1$	$q_2$	$\dots$	$q_j$	$\dots$

### 3.2.2 Абсолютно неперервний випадковий вектор

**Definition 3.2.2** Випадковий вектор  $\vec{\xi}$  називається **неперервний**, якщо

$$\xi_i, i = \overline{1, n} - \text{неперервна випадкова величина.}$$

**Proposition 3.2.3**  $\vec{\xi}$  – неперервний випадковий вектор  $\iff \forall \vec{x} \in \mathbb{R}^n : \mathbb{P}\{\vec{\xi} = \vec{x}\} = 0$ .

*TODO: довести.*

**Definition 3.2.4** Випадковий вектор  $\vec{\xi}$  називається **абсолютно неперервним**, якщо

$$\exists f_{\vec{\xi}} : F_{\vec{\xi}}(\vec{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_{\vec{\xi}}(x_1, \dots, x_n) dx_1, \dots dx_n$$

Функція  $f_{\vec{\xi}}$  називається **щільністю розподілу**.

**Proposition 3.2.5** Властивості щільності розподілу

Справедливе наступне:

- 1)  $f_{\vec{\xi}} \geq 0$ ;
- 2)  $f_{\vec{\xi}} = \frac{\partial^n F_{\vec{\xi}}}{\partial x_1 \dots \partial x_n}$ ;
- 3)  $\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x_1, \dots, x_n) dx_1, \dots dx_n = 1$ ;
- 4)  $\mathbb{P}\{\vec{\xi} \in B\} = \int_B f_{\vec{\xi}}(\vec{x}) d\vec{x}$ , якщо множина  $B$  – квадрована.

*TODO: довести.*

Обмежимося тимчасово двовимірним випадком.

**Proposition 3.2.6** Задано  $\vec{\xi} = (\xi_1, \xi_2)^T$  – абсолютно неперервний випадковий вектор (тобто існує щільність  $f_{\vec{\xi}}(x, y)$ ). Тоді існують щільності розподілу випадкових величин  $f_{\xi_1}, f_{\xi_2}$ , які називаються ще **маргинальними щільностями**. Причому

$$f_{\xi_1}(x) = \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, y) dy;$$

$$f_{\xi_2}(y) = \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, y) dx.$$

**Proof.**

$$\text{З одного боку, } \mathbb{P}\{\xi_1 \in C\} = \mathbb{P}\{(\xi_1, \xi_2)^T \in C \times \mathbb{R}\} = \int_C \int_{\mathbb{R}} f_{\vec{\xi}}(x, y) dy dx.$$

$$\text{З іншого боку, } \mathbb{P}\{\xi_1 \in C\} = \int_C f_{\xi_1}(x) dx. \quad \blacksquare$$

### 3.3 Числові характеристики випадкових векторів

#### 3.3.1 Математичне сподівання

**Definition 3.3.1** Математичним сподіванням випадкового вектора  $\vec{\xi} = (\xi_1, \dots, \xi_n)^T$  називають вектор

$$(\mathbb{E}\xi_1, \dots, \mathbb{E}\xi_n)^T$$

Інколи ще це називають **центром розсіювання**.

Зупинимося для двовимірного випадкового вектора. Для дискретної випадкової величини ми рахуємо математичне сподівання як  $\mathbb{E}\xi_1 = \sum_i x_i p_i$ . Із іншого боку, коли ми розмовляли про маргінальні ряди розподілу, ми з'ясували, що  $p_i = \sum_j p_{ij}$ . Отже, покоординатно математичне сподівання в дискретному випадку можна рахувати таким чином:

$$\mathbb{E}\xi_1 = \sum_{i,j} x_i p_{ij}$$

$$\mathbb{E}\xi_2 = \sum_{i,j} y_j p_{ij}$$

Для абсолютно неперервної випадкової величини ми рахуємо математичне сподівання як  $\mathbb{E}\xi_1 = \int_{-\infty}^{+\infty} x f_{\xi_1}(x) dx$ . Знову ж таки, ми щось вже знаємо про маргінальну щільність розподілу, як-от  $f_{\xi_1}(x) = \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, y) dy$ . Отже, покоординатно математичне сподівання в дискретному випадку можна рахувати таким чином:

$$\mathbb{E}\xi_1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{\vec{\xi}}(x, y) dx$$

$$\mathbb{E}\xi_2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{\vec{\xi}}(x, y) dy$$

#### 3.3.2 Коваріація (кореляційний момент)

**Definition 3.3.2** Коваріацією випадкового вектора  $\vec{\xi} = (\xi_1, \xi_2)^T$  називають таку величину:

$$\text{cov}(\xi_1, \xi_2) = \mathbb{E}(\xi_1 - \mathbb{E}\xi_1)(\xi_2 - \mathbb{E}\xi_2)$$

Коваріація описує теж степінь розкиду значень  $\xi_1, \xi_2$  навколо  $(\mathbb{E}\xi_1, \mathbb{E}\xi_2)^T$ .

Розпишемо детально формулу коваріації:

$$\text{cov}(\xi_1, \xi_2) = \mathbb{E}(\xi_1 \xi_2 - \xi_1 \mathbb{E}\xi_2 - \xi_2 \mathbb{E}\xi_1 + \mathbb{E}\xi_1 \mathbb{E}\xi_2) = \mathbb{E}(\xi_1 \xi_2) - \mathbb{E}\xi_2 \mathbb{E}\xi_1 - \mathbb{E}\xi_1 \mathbb{E}\xi_2 + \mathbb{E}\xi_1 \mathbb{E}\xi_2.$$

Таким чином, коваріацію можна обчислювати більш зручним чином:

$$\text{cov}(\xi_1, \xi_2) = \mathbb{E}(\xi_1 \xi_2) - \mathbb{E}\xi_1 \mathbb{E}\xi_2$$

Тільки варто зрозуміти, яким чином треба рахувати математичне сподівання  $\mathbb{E}(\xi_1 \xi_2)$ . У залежності від типу випадкової величини (дискретна чи абсолютно неперервна) маємо таке:



$$\mathbb{E}(\xi_1 \xi_2) = \sum_{i,j} (x_i y_j) p_{ij};$$

$$\mathbb{E}(\xi_1 \xi_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (xy) f_{\vec{\xi}}(x, y) dx dy.$$

Чому саме такі формули, це буде моїм боржком. Доведемо скоро.

### Proposition 3.3.3 Властивості коваріації

Справдливе наступне:

- 1)  $\text{cov}(\xi_1, \xi_1) = \mathbb{D}\xi_1$ ;
- 2) Нехай  $\xi_1, \xi_2$  – незалежні випадкові величини, тоді  $\text{cov}(\xi_1, \xi_2) = 0$ ;
- 3)  $\text{cov}(\xi_2, \xi_1) = \text{cov}(\xi_2, \xi_1)$ ;
- 4)  $\text{cov}(\xi, c) = 0$ ;
- 5)  $\text{cov}(\alpha \xi'_1 + \beta \xi''_1, \xi_2) = \alpha \text{cov}(\xi'_1, \xi_2) + \beta \text{cov}(\xi''_1, \xi_2)$ ;
- 6) Нехай  $\xi_1, \xi_2$  не є незалежними випадковими величинами, тоді  $\mathbb{D}(\xi_1 + \xi_2) = \mathbb{D}\xi_1 + \mathbb{D}\xi_2 + 2 \text{cov}(\xi_1, \xi_2)$  (узагальнення формули дисперсії суми);
- 7)  $|\text{cov}(\xi_1, \xi_2)| \leq \sqrt{\mathbb{D}\xi_1 \mathbb{D}\xi_2}$ .

*Вправа: довести. Останню властивість доведемо в наступному розділі.*

**Remark 3.3.4** Властивості 3), 5) свідчать про те, що коваріація є білінійною симетричною формою.

**Definition 3.3.5** Випадкові величини  $\xi_1, \xi_2$  називаються **некорельованими**, якщо

$$\text{cov}(\xi_1, \xi_2) = 0$$

**Remark 3.3.6**  $\xi_1, \xi_2$  – незалежні  $\implies \xi_1, \xi_2$  – некорельовані (властивість 2)).

У зворотний бік це не працює (приклад пізніше).

## 3.4 Коваріаційна матриця

**Definition 3.4.1** Коваріаційною матрицею випадкового вектора  $\vec{x} = (\xi_1, \dots, \xi_n)^T$  називають таку матрицю:

$$C_{\vec{\xi}} = \mathbb{E}(\vec{\xi} - \mathbb{E}\vec{\xi})(\vec{\xi} - \mathbb{E}\vec{\xi})^T$$

Якщо розписати розгорнуто, то отримаємо такий вигляд:

$$C_{\vec{\xi}} = \begin{pmatrix} \mathbb{D}\xi_1 & \text{cov}(\xi_1, \xi_2) & \dots & \text{cov}(\xi_1, \xi_n) \\ \text{cov}(\xi_2, \xi_1) & \mathbb{D}\xi_2 & \dots & \text{cov}(\xi_2, \xi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(\xi_n, \xi_1) & \text{cov}(\xi_n, \xi_2) & \dots & \mathbb{D}\xi_n \end{pmatrix}$$

**Proposition 3.4.2**  $C_{\vec{\xi}}$  – симетрична та невід’ємноозначена матриця.

**Proof.**

Симетричність випливає зі симетричності коваріації. Щодо невід’ємноозначеності матриці маємо:  $\vec{x}^T C_{\vec{\xi}} \vec{x} = \vec{x}^T \mathbb{E}(\vec{\xi} - \mathbb{E}\vec{\xi})(\vec{\xi} - \mathbb{E}\vec{\xi})^T \vec{x}$  ■

Якщо взяти коваріаційну матрицю  $2 \times 2$ , то, використавши критерій Силвестра, отримаємо властивість 7) коваріації.

**Theorem 3.4.3** Коваріаційна матриця  $C_{\vec{\xi}}$  – вироджена  $\iff$  між  $\xi_1, \dots, \xi_n$  афінна залежність.

**Proof.**

$\Rightarrow$  Дано:  $C_{\vec{\xi}}$  – вироджена, тобто  $\det C_{\vec{\xi}} = 0$ . Тобто існує вектор  $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ , для якого  $(C_{\vec{\xi}} \vec{x}, \vec{x}) = 0$ . (TODO: довести). ■

## 3.5 Рівномірний розподіл на площині

Рівномірний розподіл можна описати такою щільністю розподілу:

$$f_{\vec{\xi}}(x, y) = \begin{cases} \frac{1}{S(A)}, & (x, y) \in A \\ 0, & (x, y) \notin A \end{cases}$$

Такий розподіл можна ось так коротко записати:  $\xi \in U(A)$ .

## 4 Функції від випадкових величин та векторів

### 4.1 Одновимірна функція

Нехай  $\xi$  – випадкова величина. Розглянемо нову випадкову величину  $\eta = \varphi(\xi)$ , де  $\varphi$  – деяка вимірна функція.

#### 4.1.1 Дискретна випадкова величина

Маємо ряд розподілу для випадкової величини  $\xi$ .

$\xi$	$x_1$	$x_2$	$\dots$	$x_n$	$\dots$
$\mathbb{P}$	$p_1$	$p_2$	$\dots$	$p_n$	$\dots$

Цілком зрозуміло, що  $\eta$  також стане дискретною випадковою величиною, що приймає значення  $y_k = \varphi(x_k)$ . Чому дорівнює  $\mathbb{P}\{\eta = y_k\}$  залежить від ситуації.

**Example 4.1.1** Розглянемо дискретну випадкову величину  $\xi$  з таким рядом розподілом:

$\xi$	1	2	3
$\mathbb{P}$	0.5	0.3	0.2

Задамо нову випадкову величину  $\eta = \xi^2 - 4\xi$ . Ряд розподілу виглядає ось так:

$\eta$	-4	-3
$\mathbb{P}$	0.3	0.7

$$\begin{aligned}\mathbb{P}\{\eta = -4\} &= \mathbb{P}\{\xi = 2\} = 0.3; \\ \mathbb{P}\{\eta = -3\} &= \mathbb{P}\{\xi = 1\} + \mathbb{P}\{\xi = 3\} = 0.5 + 0.2 = 0.7.\end{aligned}$$

#### 4.1.2 Абсолютно неперервна випадкова величина

**Example 4.1.2** Нехай  $\xi \sim \text{Exp}(\lambda)$ . Зафіксуємо випадкову величину  $\eta = [\xi]$  (ціла частина). Так ось, виявляється,  $\eta$  буде дискретною випадковою величиною, причому не абиякою.

$\mathbb{P}\{\eta = k\} = \mathbb{P}\{[\xi] = k\} = \mathbb{P}\{\xi \in [k, k+1)\} = F_\xi(k+1) - F_\xi(k) = 1 - e^{-\lambda(k+1)} - 1 + e^{-\lambda k} = (e^{-\lambda})^k (1 - e^{-\lambda})$ . Позначимо  $p = 1 - e^{-\lambda}$  та  $q = e^{-\lambda} = 1 - p$ . Тоді звідси  $\mathbb{P}\{\eta = k\} = q^k p$ .

Отже, можемо побачити, що  $\eta \sim \text{Geom}_0(1 - e^{-\lambda})$ .

**Theorem 4.1.3** Нехай  $\varphi \in C'$  – взаємно однозначна, монотонна та її обернена  $\psi \in C'$ . Тоді  $f_\eta(y) = f_\xi(\psi(y)) \cdot |\psi'(y)| \cdot \mathbb{1}_{E_\varphi}(y)$ .

**Proof.**

Зафіксуємо деяку вимірну множину  $B$ . Із одного боку,

$$\mathbb{P}\{\eta \in B\} = \int_B f_\eta(y) dy.$$

Із іншого боку, згадавши, що  $\eta = \varphi(\xi)$ , маємо

$$\mathbb{P}\{\eta \in B\} = \mathbb{P}\{\xi \in \varphi^{-1}(B)\} = \int_{\varphi^{-1}(B)} f_\xi(x) dx \quad \square$$

Ми проведемо заміну  $y = \varphi(x)$ , тобто  $x = \psi(y)$ . Звідси  $dx = |\psi'(y)| dy$ .

$$\square \int_{B \cap E_\varphi} f_\xi(\psi(y)) |\psi'(y)| dy \quad \blacksquare$$

**Example 4.1.4** Маємо випадкову величину  $\xi \sim U(0, 1)$ . З'ясувати, як розподілена випадкова величина  $\eta = -\frac{1}{\lambda} \ln \xi$ , де  $\lambda > 0$ .

Розглянемо функцію  $\varphi(x) = -\frac{\ln x}{\lambda}$ , яка є монотонною, область значень  $E_\varphi = \mathbb{R}$ . Якщо знайти обернену функцію, отримаємо  $\psi(y) = e^{-\lambda y}$ . Похідна за модулем  $|\psi'(y)| = \lambda e^{-\lambda y}$ . Таким чином, за щойно доведеною теоремою, щільність розподілу випадкової величини  $\eta$  виглядає ось так:

$$f_\eta(y) = f_\xi(\psi(y)) |\psi'(y)| \mathbb{1}_{E_\varphi}(y).$$

По-перше,  $\mathbb{1}_{E_\varphi}(y) = 1$ , тут цілком ясно. По-друге, у нас щільність  $f_\xi(x) = \begin{cases} 1, & x \in [0, 1], \\ 0, & \text{інакше} \end{cases}$ . Відпо-

відно,  $f_\xi(e^{-\lambda y}) = \begin{cases} 1, & e^{-\lambda y} \in [0, 1] \\ 0, & \text{інакше} \end{cases}$ . Зауважимо, що  $0 \leq e^{-\lambda y} \leq 1 \implies -\lambda y \leq 0 \implies y \geq 0$ .

Отже, коли  $y \geq 0$ , то ми отримаємо  $f_\eta(y) = 1 \cdot \lambda e^{-\lambda y}$ . Коли  $y < 0$ , то ми отримаємо  $f_\eta(y) = 0 \cdot \lambda e^{-\lambda y}$ . Разом отримаємо відому щільність:

$$f_\eta(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0 \\ 0, & \text{інакше} \end{cases} \implies \eta \sim \text{Exp}(\lambda).$$

**Theorem 4.1.5** Нехай  $\varphi$  розпадається на  $n$  функцій вигляду  $\varphi_i$ , які задовольняють першій теоремі (неперервні, монотонні та взаємно однозначні). Тоді  $f_\eta(y) = \sum_{i=1}^n f_{\eta_i}(y)$ , де кожна  $\eta_i = \varphi_i(\xi)$ .

*Майже аналогічне доведення.*

**Example 4.1.6** Маємо випадкову величину  $\xi \sim U(-2, 3)$ . З'ясувати, як розподілена випадкова величина  $\eta = 1 - |\xi - 2|$ .

$$\text{Одразу запишу для зручності щільність } f_\xi(x) = \begin{cases} \frac{1}{5}, & x \in [-2, 3] \\ 0, & \text{інакше} \end{cases}.$$

Розглянемо функцію  $\varphi(x) = 1 - |x - 2| = \begin{cases} x - 1, & x < 2 \\ 3 - x, & x \geq 2 \end{cases}$ . Позначу першу функцію  $\varphi_1(x) = x - 1$  та  $\varphi_2(x) = 3 - x$ . Вони обидва взаємно однозначні на своїй області визначення, монотонні. Розберемося з кожною функцією окремо.

При функції  $\varphi_1(x) = x - 1, x < 2$  маємо  $E_{\varphi_1} = (-\infty, 1]$ , обернена функція  $\psi_1(y) = y + 1$  та модуль похідної  $|\psi'_1(y)| = 1$ . Якщо позначити  $\eta_1 = \varphi_1(\xi)$ , то отримаємо  $f_{\eta_1}(y) = f_\xi(\psi_1(y))|\psi'_1(y)|\mathbb{1}_{E_{\varphi_1}}(y)$ .

$$\text{Зауважимо, що } f_\xi(\psi_1(y)) = f_\xi(y + 1) = \begin{cases} \frac{1}{5}, & y + 1 \in [-2, 3] \\ 0, & \text{інакше} \end{cases} = \begin{cases} \frac{1}{5}, & y \in [-3, 2] \\ 0, & \text{інакше} \end{cases} = \frac{1}{5} \mathbb{1}_{y \in [-3, 2]}(y).$$

$$\text{Разом отримаємо } f_{\eta_1}(y) = \frac{1}{5} \mathbb{1}_{[-3, 2]}(y) \mathbb{1}_{(-\infty, 1]}(y) = \frac{1}{5} \mathbb{1}_{[-3, 1]}(y).$$

При функції  $\varphi_2(x) = 3 - x, x \geq 2$  все цілком аналогічно робиться. Якщо позначити  $\eta_2 = \varphi_2(\xi)$ , то отримаємо  $f_{\eta_2}(y) = \frac{1}{5} \mathbb{1}_{[0, 1]}(y)$ .

$$\text{Разом отримаємо } f_\eta(y) = f_{\eta_1}(y) + f_{\eta_2}(y) = \frac{1}{5} \mathbb{1}_{[-3, 1]}(y) + \frac{1}{5} \mathbb{1}_{[0, 1]}(y) = \begin{cases} \frac{1}{5}, & y \in [-3, 0) \\ \frac{2}{5}, & y \in [0, 1] \\ 0, & \text{інакше} \end{cases}.$$

## 4.2 Багатовимірна функція

Нехай  $\vec{\xi}$  – випадковий вектор. Розглянемо новий випадковий вектор  $\eta = \varphi(\vec{\xi})$ , де  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$  – деяка вимірна функція багатьох змінних.

### 4.2.1 Дискретний випадковий вектор

Маємо випадковий вектор  $\vec{\xi}$ , який дискретний. Кожний  $\xi_i$  набуває зліченного (чи скінченного) числа значень. Тоді зокрема й  $\eta = \varphi(\xi_1, \dots, \xi_n)$  набуватиме зліченного (чи скінченного) числа значень. Отже,  $\eta$  – автоматично дискретна випадкова величина.

**Example 4.2.1** Розглянемо дискретний випадковий вектор  $\vec{\xi} = (\xi_1, \xi_2)^T$  з таблицею розподілу:

$\xi_1/\xi_2$	-1	1
1	0.1	0.2
2	0.3	0.2
3	0.1	0.1

Задамо нову випадкову величину  $\eta = \xi_1 + \xi_2$ . Ряд розподілу виглядає ось так:

$\eta$	0	1	2	3	4
$\mathbb{P}$	0.1	0.3	0.3	0.2	0.1

Наведу кілька прикладів, як ці ймовірності ми отримали.

$$\mathbb{P}\{\eta = 0\} = \mathbb{P}\{(\xi_1, \xi_2) = (1, -1)\}.$$

$$\mathbb{P}\{\eta = 2\} = \mathbb{P}\{(\xi_1, \xi_2) = (1, 1)\} + \mathbb{P}\{(\xi_1, \xi_2) = (3, -1)\} = 0.2 + 0.1 = 0.3.$$

#### 4.2.2 Абсолютно неперервний випадковий вектор

Ситуація тут набагато складніше, але викрутитися можна. Тут будуть теореми, які, насправді, дозволяють знайти випадковий вектор  $\vec{\eta} = \vec{\varphi}(\vec{\xi})$ , але це допоможе нам для нашої основної мети.

**Theorem 4.2.2** Нехай  $\vec{\varphi} \in C'$  – взаємно однозначна та її обернена  $\vec{\psi} \in C'$ . Тоді  $f_{\vec{\eta}}(\vec{y}) = f_{\vec{\xi}}(\vec{\psi}(\vec{y})) |J_{\vec{\psi}}(\vec{y})| \mathbb{1}_{E_{\vec{\varphi}}}(\vec{y})$ . Причому якобіан ненульовий.

**Theorem 4.2.3** Нехай  $\vec{\varphi}$  розпадається на  $n$  вектор-функцій вигляду  $\vec{\varphi}_i$ , які задовольняють першій теоремі (неперервні та взаємно однозначні). Тоді  $f_{\vec{\eta}}(\vec{y}) = \sum_{i=1}^n f_{\vec{\eta}_i}(\vec{y})$ , де кожний  $\vec{\eta}_i = \vec{\varphi}_i(\vec{\xi})$ .

**Example 4.2.4** Маємо довільний випадковий вектор  $\vec{\xi} = (\xi_1, \xi_2)^T$ , який абсолютно неперервний. З'ясувати, чи буде випадкова величина  $\eta = \xi_1 + \xi_2$  абсолютно неперервною. Якщо так, знайти щільність.

Тут треба зробити деякий трюк. Ми розглянемо випадковий вектор  $\vec{\eta} = (\xi_1 + \xi_2, \xi_1 - \xi_2)^T$ , де друга координата – будь-яка допоміжна, але така, щоб було з нею зручно працювати. Ми маємо вектор-функцію  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \xrightarrow{\vec{\varphi}} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}$ . Зауважимо, що  $E_{\vec{\varphi}} = \mathbb{R}^2$ . Далі, оборотна функція

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \xrightarrow{\vec{\psi}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(y_1 + y_2) \\ \frac{1}{2}(y_1 - y_2) \end{pmatrix}, \text{ а якобіан } J_{\vec{\psi}}(\vec{y}) = \det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = -\frac{1}{2}.$$

$$f_{\vec{\eta}}(\vec{y}) = f_{\vec{\xi}}(\vec{\psi}(\vec{y})) |J_{\vec{\psi}}(\vec{y})| \mathbb{1}_{E_{\vec{\varphi}}}(\vec{y}) = \frac{1}{2} f_{\vec{\xi}}(\vec{\psi}(\vec{y})).$$

Із іншого боку,  $f_{\vec{\eta}}(\vec{y}) = f_{\xi_1 + \xi_2, \xi_1 - \xi_2}(y_1, y_2) = \frac{1}{2} f_{\vec{\xi}}\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right)$ . Нам треба щільність розподілу лише від  $\xi_1 + \xi_2 = \eta$ , тобто треба маргінальну щільність. Це можна знайти таким чином:

$$\begin{aligned} f_{\eta}(y_1) &= f_{\xi_1 + \xi_2}(y_1) = \int_{-\infty}^{+\infty} f_{\xi_1 + \xi_2, \xi_1 - \xi_2}(y_1, y_2) dy_2 = \frac{1}{2} \int_{-\infty}^{+\infty} f_{\vec{\xi}}\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right) \frac{y_1 + y_2}{2} = x \\ &= \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, y_1 - x) dx. \end{aligned}$$

Будемо цю формулу часто застосовувати, тому наведу окреме твердження.

**Proposition 4.2.5** Маємо довільний випадковий вектор  $\vec{\xi} = (\xi_1, \xi_2)^T$ , який абсолютно неперервний. Тоді  $\eta = \xi_1 + \xi_2$  буде абсолютно неперервною також, а її щільність

$$f_{\eta}(y) = \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, y - x) dx.$$

Щоправда, так не завжди зручно рахувати, тому розглянемо другий спосіб. Ми краще знайдемо спочатку функцію розподілу  $F_{\eta}(z) = \mathbb{P}\{\eta \leq z\}$ . Якщо розписати ймовірність, буде  $\mathbb{P}\{\varphi(\xi_1, \xi_2) \leq z\}$ , а всередині ймовірності записана множина  $\{(x, y) \in \mathbb{R}^2 : \varphi(x, y) \leq z\} \stackrel{\text{позн.}}{=} D_z$ . Звідси випливає, що

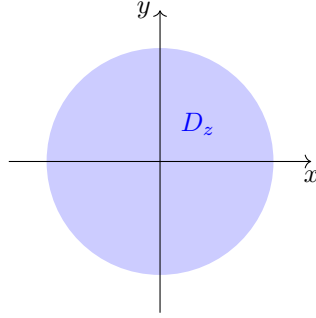
$$F_{\eta}(z) = \mathbb{P}\{\vec{\xi} \in D_z\} = \iint_{D_z} f_{\vec{\xi}}(x, y) dx dy.$$

**Example 4.2.6** Розглянемо випадкові величини  $\xi_1, \xi_2 \sim \mathcal{N}(0, 1)$ , причому вони незалежні. Визначимо випадкову величину  $\eta = \xi_1^2 + \xi_2^2$  та з'ясуємо, як вона розподілена.

Користуватися будемо другим способом, маємо  $F_{\eta}(z) = \mathbb{P}\{\eta \leq z\} = \mathbb{P}\{\xi_1^2 + \xi_2^2 \leq z\}$ .

Якщо  $z < 0$ , то цілком зрозуміло, що  $\mathbb{P}\{\xi_1^2 + \xi_2^2 \leq z\} = 0$ , тож звідси  $F_{\eta}(z) = 0$ .

Якщо  $z \geq 0$ , то позначимо область  $D_z = \{(x, y) : x^2 + y^2 \leq z\}$



$D_z$  описує коло з центром на початку координат та радіуса  $R = \sqrt{z}$ .

$$\begin{aligned} \text{Тоді ми маємо } F_\eta(z) &= \mathbb{P}\{\vec{\xi} \in D_z\} = \iint_{D_z} f_{\vec{\xi}}(x, y) dx dy \stackrel{\xi_1, \xi_2 \text{ незалежні}}{=} \iint_{D_z} f_{\xi_1}(x) f_{\xi_2}(y) dx dy \stackrel{\text{полярна заміна}}{=} \\ &= \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^{\sqrt{z}} \rho e^{-\frac{\rho^2}{2}} d\rho = \dots = 1 - e^{-\frac{z}{2}}. \end{aligned}$$

Отже, ми разом отримали  $F_\eta(z) = \begin{cases} 1 - e^{-\frac{z}{2}}, & z \geq 0 \\ 0, & \text{інакше} \end{cases}$ . Продиференціюємо дану функцію розподілу

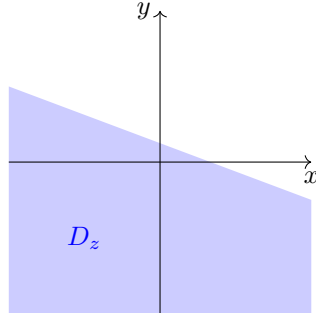
$$\text{— отримаємо щільність } f_\eta(z) = \begin{cases} \frac{1}{2} e^{-\frac{z}{2}}, & z \geq 0 \\ 0, & \text{інакше} \end{cases} \implies \eta \sim \text{Exp}\left(\frac{1}{2}\right).$$

### 4.3 Щільності розподілу суми, добутку та частки випадкової величини

Маємо випадковий вектор  $\vec{\xi} = (\xi_1, \xi_2)^T$ , який асбсолютно неперервний. Задача полягає в тому, щоб знайти щільності  $f_{\xi_1 + \xi_2}$ ,  $f_{\xi_1 \cdot \xi_2}$ ,  $f_{\frac{\xi_2}{\xi_1}}$ .

#### 4.3.1 $\xi_1 + \xi_2$

$$F_{\xi_1 + \xi_2}(z) = \mathbb{P}\{\xi_1 + \xi_2 \leq z\} =$$

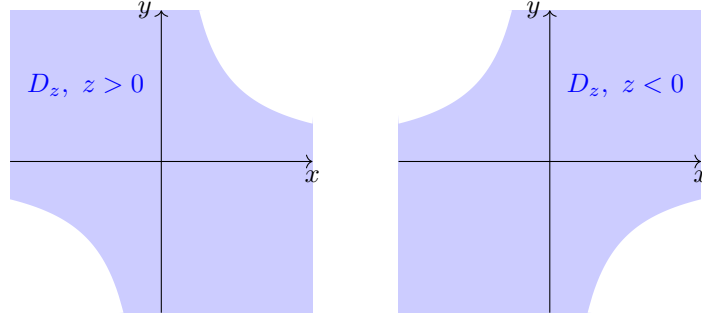


$D_z$  описується нерівністю  $x + y \leq z$ . Пряма перетинає вісь  $OY$  на висоті  $z$ .

$$\begin{aligned} &= \mathbb{P}\{\vec{\xi} \in D_z\} = \iint_{D_z} f_{\vec{\xi}}(x, y) dx dy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f_{\vec{\xi}}(x, y) dy. \\ f_{\xi_1 + \xi_2}(z) &= \frac{d}{dz} \left( \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f_{\vec{\xi}}(x, y) dy \right) = \int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, z-x) dx. \\ \text{Якщо } \xi_1, \xi_2 \text{ незалежні, то звідси } f_{\xi_1 + \xi_2}(z) &= \int_{-\infty}^{+\infty} f_{\xi_1}(x) f_{\xi_2}(z-x) dx. \end{aligned}$$

#### 4.3.2 $\xi_1 \cdot \xi_2$

$$F_{\xi_1 \cdot \xi_2}(z) = \mathbb{P}\{\xi_1 \cdot \xi_2 \leq z\} =$$



$D_z$  описується нерівністю  $x \cdot y \leq z$ . Два малюнки в залежності від  $z$ .

при  $z > 0$  (можна навіть включити  $z = 0$ )  $\boxed{=}$

$$f_{\xi_1 \xi_2}(z) = - \int_{-\infty}^0 f_{\bar{\xi}}\left(x, \frac{z}{x}\right) \frac{1}{x} dx + \int_0^{+\infty} f_{\bar{\xi}}\left(x, \frac{z}{x}\right) \frac{1}{x} dx = \int_{-\infty}^{+\infty} f_{\bar{\xi}}\left(x, \frac{z}{x}\right) \frac{1}{|x|} dx.$$

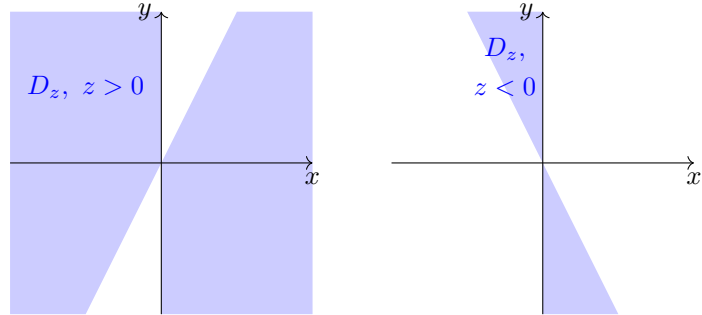
при  $z < 0$   $\boxed{=}$

$$f_{\xi_1 \xi_2}(z) \stackrel{\text{отримаємо аналогічно}}{=} \dots = \int_{-\infty}^{+\infty} f_{\bar{\xi}}\left(x, \frac{z}{x}\right) \frac{1}{|x|} dx.$$

Якщо  $\xi_1, \xi_2$  незалежні, то звідси  $f_{\xi_1 \xi_2}(z) = \int_{-\infty}^{+\infty} f_{\xi_1}(x) f_{\xi_2}\left(\frac{z}{x}\right) \frac{1}{|x|} dx$ .

#### 4.3.3 $\frac{\xi_2}{\xi_1}$

$$F_{\frac{\xi_2}{\xi_1}}(z) = \mathbb{P}\left\{\frac{\xi_2}{\xi_1} \leq z\right\} =$$



$D_z$  описується нерівністю  $\frac{y}{x} \leq z$ . Два малюнки в залежності від  $z$ .

при  $z > 0$  (можна навіть включити  $z = 0$ )  $\boxed{=}$

$$f_{\frac{\xi_2}{\xi_1}}(z) = - \int_{-\infty}^0 f_{\bar{\xi}}(x, zx) x dx + \int_0^{+\infty} f_{\bar{\xi}}(x, zx) x dx = \int_{-\infty}^{+\infty} f_{\bar{\xi}}(x, zx) |x| dx.$$

Аналогічно треба окремо перевіряти при  $z < 0$ , але формула буде такою самою.

Якщо  $\xi_1, \xi_2$  незалежні, то звідси  $f_{\frac{\xi_2}{\xi_1}}(z) = \int_{-\infty}^{+\infty} f_{\xi_1}(x) f_{\xi_2}(zx) |x| dx$ .

**Example 4.3.1** Розглянемо випадкові величини  $\xi_1, \xi_2 \in \mathcal{N}(0, 1)$ , причому вони незалежні. Знайти щільність розподілу випадкової величини  $\eta = \frac{\xi_2}{\xi_1}$ .

Мабуть, одразу застосую готову формулу, бо перевиводити її сенсу жодного.

$$f_{\eta}(z) = \int_{-\infty}^{+\infty} f_{\xi_1}(x) f_{\xi_2}(zx) |x| dx = \int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{-\frac{x^2}{2}} e^{-\frac{(zx)^2}{2}} |x| dx = \frac{1}{\pi} \int_0^{+\infty} e^{-\frac{x^2}{2} - \frac{z^2 x^2}{2}} x dx =$$

$$= \frac{1}{\pi} \frac{1}{z^2 + 1} \int_0^{+\infty} e^{-\frac{z^2 + 1}{2} x^2} d\left(\frac{z^2 + 1}{2} x^2\right) = \frac{1}{\pi} \frac{1}{z^2 + 1} \int_0^{+\infty} e^{-t} dt = \frac{1}{\pi} \frac{1}{z^2 + 1}.$$

До речі, випадкова величина, що має такий розподіл  $f_{\eta}(z) = \frac{1}{\pi} \frac{1}{z^2 + 1}$ , називається **розподілом Коші** та позначають коротко  $\eta \sim \text{Cauchy}$ .

#### 4.4 Розподіл максимуму та мінімуму

Нехай  $\xi_1, \dots, \xi_n$  – незалежні (!) випадкові величини. Задамо випадкову величину  $M = \max\{\xi_1, \dots, \xi_n\}$  та знайдемо функцію розподілу:

$$F_M(x) = \mathbb{P}\{M \leq x\} = \mathbb{P}\{\max\{\xi_1, \dots, \xi_n\} \leq x\} = \mathbb{P}\{\xi_1 \leq x, \dots, \xi_n \leq x\} = \mathbb{P}\{\xi_1 \leq x\} \dots \mathbb{P}\{\xi_n \leq x\} = F_{\xi_1}(x) \dots F_{\xi_n}(x).$$

$$\text{Отже, } F_M(x) = \prod_{i=1}^n F_{\xi_i}(x).$$

Якщо ще додатково попросити, щоб  $\xi_1, \dots, \xi_n$  були однаково розподіленими абсолютно неперервними, то  $F_M(x) = F^n(x)$ . Внаслідок чого отримаємо щільність:

$$f_M(x) = n F^{n-1}(x) f(x).$$

Нехай  $\xi_1, \dots, \xi_n$  – незалежні (!) випадкові величини. Задамо випадкову величину  $m = \min\{\xi_1, \dots, \xi_n\}$  та знайдемо функцію розподілу:

$$F_m(x) = \mathbb{P}\{m \leq x\} = \mathbb{P}\{\min\{\xi_1, \dots, \xi_n\} \leq x\} = 1 - \mathbb{P}\{\min\{\xi_1, \dots, \xi_n\} > x\} =$$

$$= 1 - \mathbb{P}\{\xi_1 > x, \dots, \xi_n > x\} = 1 - \mathbb{P}\{\xi_1 > x\} \dots \mathbb{P}\{\xi_n > x\} = 1 - (1 - \mathbb{P}\{\xi_1 \leq x\}) \dots (1 - \mathbb{P}\{\xi_n \leq x\}) =$$

$$= 1 - (1 - F_{\xi_1}(x)) \dots (1 - F_{\xi_n}(x)).$$

$$\text{Отже, } F_m(x) = 1 - \prod_{i=1}^n (1 - F_{\xi_i}(x)).$$

Якщо ще додатково попросити, щоб  $\xi_1, \dots, \xi_n$  були однаково розподіленими абсолютно неперервними, то  $F_m(x) = 1 - (1 - F(x))^n$ . Внаслідок чого отримаємо щільність:

$$f_m(x) = n(1 - F(x))^{n-1} f(x).$$

#### 4.5 Математичне сподівання функцій від випадкових величин/векторів

У нас була формула математичного сподівання для випадкової величини  $\xi$ . Це було таке:

$$\mathbb{E}\xi = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx.$$

Колись ми рахували дисперсію та я казав, що  $\mathbb{E}\xi^2 = \int_{-\infty}^{+\infty} x^2 f_{\xi}(x) dx$ . Час відповісти на питання, чому ми допускали таку рівність.

Нехай  $\xi$  – дискретна випадкова величина та  $\eta = \varphi(\xi)$  (яка автоматично дискретна). Маємо  $E\varphi = \{y_1, y_2, \dots\}$  (може бути скінченна чи зліченна множина). Кожне з цих значень  $y_i$  приймається хоча б на одній з точках  $\{x_1, x_2, \dots\}$ . Наприклад, нехай  $y_i = \varphi(x)$  при  $x \in \{x_{s_1}, x_{s_2}, \dots\}$  (знову ж таки, може бути скінченне чи зліченне число значень, але точно одна точка  $\epsilon$ ). Тоді

$$y_i \mathbb{P}\{\eta = y_i\} = \sum_k y_i \mathbb{P}\{\xi = x_{s_k}\} = \sum_k \varphi(x_{s_k}) \mathbb{P}\{\xi = x_{s_k}\}.$$

Тоді після перегрупування ми отримаємо  $\mathbb{E}\varphi(\xi) = \sum_i \varphi(x_i) p_i$ . (TODO: а збіжність?)

**Theorem 4.5.1** Нехай заданий випадковий вектор  $\vec{\xi} = (\xi_1, \dots, \xi_n)^T$  та вимірна функція  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ .

$$\text{Тоді } \mathbb{E}\varphi(\xi) = \int \dots \int_{\mathbb{R}^n} \varphi(x_1, \dots, x_n) f_{\vec{\xi}}(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Напередодні доведення наведемо корисну лему.

**Lemma 4.5.2** Нехай  $\mathbb{P}\{\eta \geq 0\} = 1$ . Тоді  $\mathbb{E}\eta = \int_0^{+\infty} (1 - F_{\eta}(x)) dx$ .

**Proof.**

Зауважимо, що вираз  $1 - F_\eta(x) = \mathbb{P}\{\eta > x\} = \int_x^{+\infty} f_\eta(z) dz$ . Тоді звідси

$$\int_0^{+\infty} (1 - F_\eta(x)) dx = \int_0^{+\infty} \int_x^{+\infty} f_\eta(z) dz dx = \int_0^{+\infty} \int_0^z f_\eta(z) dx dz = \int_0^{+\infty} f_\eta(z) z dz = \mathbb{E}\eta.$$

Оскільки  $\eta$  невід'ємна майже скрізь, то звідси  $f_\eta(z) = 0$  при  $z < 0$ . Все це ми допускали, коли  $\eta$  – абсолютно неперервна випадкова величина. ■

Повертаємося до теореми.

**Proof.**

Доведення проведемо для функції двох змінних.

I.  $\varphi$  – невід'ємна.

Уведемо множину  $D_z = \{(x, y) \in \mathbb{R}^2 : \varphi(x, y) \geq z\}$ . Тоді матимемо:

$$\begin{aligned} \mathbb{E}\varphi &= \mathbb{E}\varphi(\xi_1, \xi_2) \stackrel{\text{Лм. 4.5.2}}{=} \int_0^{+\infty} \mathbb{P}\{\varphi(\xi_1, \xi_2) \geq z\} dz = \int_0^{+\infty} \mathbb{P}\{\vec{\xi} \in D_z\} dz = \int_0^{+\infty} \left( \iint_{D_z} f_{\vec{\xi}}(x, y) dx dy \right) dz = \\ &= \iint_{\mathbb{R}^2} \left( \int_0^{\varphi(x, y)} f_{\vec{\xi}}(x, y) dz \right) dx dy = \iint_{\mathbb{R}^2} f_{\vec{\xi}}(x, y) \varphi(x, y) dx dy. \end{aligned}$$

II.  $\varphi$  – довільна.

Запишемо функцію  $\varphi(x, y) = \varphi_+(x, y) - \varphi_-(x, y)$ , де

$$\varphi_+(x, y) = \max\{\varphi(x, y), 0\} \geq 0 \quad \varphi_-(x, y) = -\min\{\varphi(x, y), 0\}.$$

$$\begin{aligned} \mathbb{E}\varphi &= \mathbb{E}\varphi(\xi_1, \xi_2) = \mathbb{E}(\varphi_+(\xi_1, \xi_2) - \varphi_-(\xi_1, \xi_2)) = \mathbb{E}\varphi_+(\xi_1, \xi_2) - \mathbb{E}\varphi_-(\xi_1, \xi_2) \stackrel{\text{пункт I}}{=} \\ &= \iint_{\mathbb{R}^2} f_{\vec{\xi}}(x, y) \varphi_+(x, y) dx dy - \iint_{\mathbb{R}^2} f_{\vec{\xi}}(x, y) \varphi_-(x, y) dx dy = \iint_{\mathbb{R}^2} f_{\vec{\xi}}(x, y) \varphi(x, y) dx dy. \end{aligned} \quad \blacksquare$$

Тепер повернімося до властивостей

$$\mathbb{E}(\xi + \eta) = \mathbb{E}\xi + \mathbb{E}\eta;$$

$$\mathbb{E}(\xi_1 \xi_2) = \mathbb{E}\xi_1 \mathbb{E}\xi_2 \quad (\text{тут для незалежних величин}).$$

Ми довели це лише для дискретного випадку. Тепер можемо довести для абсолютно неперервних.

$$\mathbb{E}(\xi + \eta) = \iint_{\mathbb{R}^2} (x + y) f_{(\xi, \eta)^T}(x, y) dx dy = \iint_{\mathbb{R}^2} x f_{(\xi, \eta)^T}(x, y) dx dy + \iint_{\mathbb{R}^2} y f_{(\xi, \eta)^T}(x, y) dx dy = \mathbb{E}\xi + \mathbb{E}\eta.$$

$$\mathbb{E}(\xi_1 \xi_2) = \iint_{\mathbb{R}^2} x_1 x_2 f_{(\xi_1, \xi_2)^T}(x_1, x_2) dx_1 dx_2 = \left( \int_{\mathbb{R}} x_1 f_{\xi_1}(x_1) dx_1 \right) \left( \int_{\mathbb{R}} x_2 f_{\xi_2}(x_2) dx_2 \right) = \mathbb{E}\xi_1 \mathbb{E}\xi_2.$$



## 5 Характеристичні функції

### 5.1 Характеристичні функції випадкових величин

**Definition 5.1.1** Характеристичною функцією випадкової величини  $\xi$  називають функцію  $\chi_\xi: \mathbb{R} \rightarrow \mathbb{C}$ , що задана такою формулою:

$$\chi_\xi(t) = \mathbb{E}(e^{it\xi})$$

У випадку дискретної випадкової величини маємо

$$\chi_\xi(t) = \sum_i e^{itx_i} p_i$$

У випадку абсолютно неперервної випадкової величини маємо

$$\chi_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} f_\xi(x) dx$$

**Remark 5.1.2** Те, що записано інтегралом, – це, насправді кажучи, перетворення Фур'є з комплексного аналізу. За допомогою оберненого перетворення Фур'є ми зможемо знайти саму щільність, знаючи характеристичну функцію.

#### Proposition 5.1.3 Властивості характеристичної функції

Справедливе наступне:

- 1)  $\chi_\xi(0) = 1$ ;
- 2)  $|\chi_\xi(t)| \leq 1$ ;
- 3)  $\chi \in C(\mathbb{R})$ ;
- 4) Якщо  $\xi_1, \xi_2$  – незалежні випадкові величини, то  $\chi_{\xi_1+\xi_2}(t) = \chi_{\xi_1}(t)\chi_{\xi_2}(t)$ ;
- 5) Якщо існує  $\mathbb{E}\xi^n$ , то  $\mathbb{E}\xi^n = \frac{1}{i^n}\chi_\xi^{(n)}(0)$ ;
- 6)  $\chi_{a\xi+b}(t) = e^{ibt}\chi_\xi(at)$ ;
- 7)  $\chi$  парна  $\iff \chi$  – дійснозначна.

**Proof.**

Доведемо всі властивості:

1)  $\chi_\xi(0) = \mathbb{E}e^{i \cdot 0 \xi} = \mathbb{E}1 = 1$ ;

2)  $|\chi_\xi(t)| = |\mathbb{E}e^{it\xi}| = |\mathbb{E} \cos t\xi + i\mathbb{E} \sin t\xi| = \sqrt{(\mathbb{E} \cos t\xi)^2 + (\mathbb{E} \sin t\xi)^2} \leq \sqrt{(\mathbb{E}\xi)^2} \leq \mathbb{E}\xi^2$ , через те, що  $\mathbb{D}\xi \geq 0$ .  
 $\leq \sqrt{\mathbb{E}(\cos t\xi)^2 + \mathbb{E}(\sin t\xi)^2} = \sqrt{\mathbb{E}(\cos^2 t\xi + \sin^2 t\xi)} = 1$ .

3) Нехай  $\xi$  – АНВВ. Тобто маємо  $\chi_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} f_\xi(x) dx$ . Зазначимо, що даний інтеграл рівномірно збіжний на  $\mathbb{R}$ . Дійсно,  $|e^{itx} f_\xi(x)| \leq f_\xi(x) = M(x)$  – мажоранта. Далі,  $\int_{-\infty}^{+\infty} M(x) dx = 1 < \infty$  – збіжний. Далі застосовуємо ознаку Ваєрштраса. Оскільки  $f_\xi$

■

## 6 Граничні теореми

### 6.1 Деякі важливі нерівності

#### Lemma 6.1.1 Нерівність Маркова

Нехай  $\eta$  – випадкова величина така, що  $\mathbb{P}\{\eta \geq 0\} = 1$ . Тоді  $\mathbb{P}\{\eta \geq a\} \leq \frac{\mathbb{E}\eta}{a}$ .

**Proof.**

$\mathbb{E}\eta = \mathbb{E}(\eta \mathbb{1}\{0 \leq \eta < a\} + \eta \mathbb{1}\{\eta \geq a\}) = \mathbb{E}\eta \mathbb{1}\{\eta < a\} + \mathbb{E}\eta \mathbb{1}\{\eta \geq a\} \geq \mathbb{E}\eta \mathbb{1}\{\eta \geq a\} \geq \mathbb{E}\{a \mathbb{1}\{\eta \geq a\}\} = a \mathbb{E}\mathbb{1}\{\eta \geq a\} = a \mathbb{P}\{\eta \geq a\}$ . ■

#### Theorem 6.1.2 Нерівність Чебишова

Нехай  $\xi$  – випадкова величина, де  $\exists \mathbb{E}\xi < \infty$ ,  $\exists \mathbb{D}\xi < \infty$ . Тоді  $\mathbb{P}\{|\xi - \mathbb{E}\xi| \geq a\} \leq \frac{\mathbb{D}\xi}{a^2}$  для будь-якого  $a > 0$ .

**Proof.**

Розглянемо випадкову величину  $\eta = (\xi - \mathbb{E}\xi)^2$ , при цьому  $\mathbb{P}\{\eta \geq 0\} = 1$ . Тоді за нерівністю Маркова  $\mathbb{P}\{\eta \geq a^2\} \leq \frac{\mathbb{E}\eta}{a^2}$ . Розгорнуто це буде так:

$$\mathbb{P}\{|\xi - \mathbb{E}\xi| \geq a\} = \mathbb{P}\{(\xi - \mathbb{E}\xi)^2 \geq a^2\} \leq \frac{\mathbb{E}(\xi - \mathbb{E}\xi)^2}{a^2} = \frac{\mathbb{D}\xi}{a^2}. \quad \blacksquare$$

### 6.2 Види збіжностей

Задамо послідовність випадкових величин  $\{\xi_n, n \geq 1\}$ .

**Definition 6.2.1** Послідовність  $\{\xi_n\}$  **збігається майже скрізь** (або **за ймовірністю 1**) до випадкової величини  $\xi$ , якщо

$$\mathbb{P}\{\xi_n(\omega) \rightarrow \xi(\omega), n \rightarrow \infty\} = 1$$

Позначення:  $\xi_n \xrightarrow{\text{м.с.}} \xi$ .

**Definition 6.2.2** Послідовність  $\{\xi_n\}$  **збігається за середнім квадратичним** до випадкової величини  $\xi$ , якщо

$$\mathbb{E}(\xi_n - \xi)^2 \rightarrow 0, n \rightarrow \infty$$

Позначення:  $\xi_n \xrightarrow{\mathbb{L}_2} \xi$ .

**Definition 6.2.3** Послідовність  $\{\xi_n\}$  **збігається за середнім** до випадкової величини  $\xi$ , якщо

$$\mathbb{E}|\xi_n - \xi| \rightarrow 0, n \rightarrow \infty$$

Позначення:  $\xi_n \xrightarrow{\mathbb{L}_1} \xi$ .

**Definition 6.2.4** Послідовність  $\{\xi_n\}$  **збігається за ймовірністю** до випадкової величини  $\xi$ , якщо

$$\forall \varepsilon > 0 : \mathbb{P}\{|\xi_n - \xi| > \varepsilon\} \rightarrow 0, n \rightarrow \infty$$

Позначення:  $\xi_n \xrightarrow{\mathbb{P}} \xi$ .

**Remark 6.2.5** Еквівалентне означення:  $\forall \varepsilon > 0 : \mathbb{P}\{|\xi_n - \xi| \leq \varepsilon\} \rightarrow 1, n \rightarrow \infty$ .

**Definition 6.2.6** Послідовність  $\{\xi_n\}$  **збігається за розподілом** до випадкової величини  $\xi$ , якщо

$$\forall x \in \mathbb{R} : F_\xi - \text{неперервна} : F_{\xi_n}(x) \rightarrow F_\xi(x), n \rightarrow \infty$$

Позначення:  $\xi_n \xrightarrow{d} \xi$ .

**Example 6.2.7** Розглянемо послідовність випадкових величин  $\{\xi_n, n \geq 1\}$ , кожна з яких має ряд розподілу:

$$\begin{array}{c|c|c|c} \xi_n & -n+1 & 1+\frac{1}{n} & n+3 \\ \hline \mathbb{P} & \frac{1}{2n} & 1-\frac{1}{n} & \frac{1}{2n} \end{array}$$

Імовіріше за всього гранична випадкова величина  $\xi = 1$ , оскільки зі збільшенням  $n$  випадкова величина  $\xi_n$  частіше набуватиме значення 1. Перевіримо на збіжність.

$$\mathbb{E}(\xi_n - \xi)^2 = \mathbb{E}(\xi_n - 1)^2 = (-n-2)^2 \cdot \frac{1}{2n} + \frac{1}{n^2} \cdot \left(1 - \frac{1}{n}\right) + (n+2)^2 \cdot \frac{1}{2n} \rightarrow +\infty \text{ при } n \rightarrow \infty.$$

$$\mathbb{E}|\xi_n - \xi| = \mathbb{E}|\xi_n - 1| = |-n-2| \cdot \frac{1}{2n} + \left|\frac{1}{n}\right| \cdot \left(1 - \frac{1}{n}\right) + |n+2| \cdot \frac{1}{2n} = \frac{n+2}{2n} + \frac{1}{n} \left(1 - \frac{1}{n}\right) + \frac{n+2}{2n} \rightarrow \frac{1}{2} + 0 + \frac{1}{2} = 1 \not\rightarrow 0 \text{ при } n \rightarrow \infty.$$

$$\mathbb{P}\{|\xi_n - \xi| > \varepsilon\} = \mathbb{P}\{\xi_n \notin [\xi - \varepsilon, \xi + \varepsilon]\} = \mathbb{P}\{\xi_n \notin [1 - \varepsilon, 1 + \varepsilon]\} = \mathbb{P}\{\xi_n = -n+1\} + \mathbb{P}\{\xi_n = n+3\} = \frac{1}{n} \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Отже, нема збіжності за  $\mathbb{L}_2$ ,  $\mathbb{L}_1$ , але є збіжність за  $\mathbb{P}$ .

**Lemma 6.2.8**  $\mathbb{E}|\eta| \leq \sqrt{\mathbb{E}\eta^2}$ .

**Proof.**

Дійсно, маємо  $\mathbb{D}|\eta| \geq 0$ . Якщо розписати дисперсію, буде  $\mathbb{E}(|\eta|^2) - (\mathbb{E}|\eta|)^2 = \mathbb{E}(\eta^2) - (\mathbb{E}|\eta|)^2 \geq 0$ . Власне, звідси й випливає  $\mathbb{E}|\eta| \leq \sqrt{\mathbb{E}\eta^2}$ . ■

**Proposition 6.2.9** Задана послідовність  $\{\xi_n, n \geq 1\}$ . Тоді маємо наступний перехід:

$$\xi_n \xrightarrow{\mathbb{L}_2} \xi \implies \xi_n \xrightarrow{\mathbb{L}_1} \xi \implies \xi_n \xrightarrow{\mathbb{P}} \xi.$$

**Proof.**

$$\mathbb{L}_2 \implies \mathbb{L}_1.$$

Нехай  $\xi_n \xrightarrow{\mathbb{L}_2} \xi$ . Тоді  $0 \leq \mathbb{E}|\xi_n - \xi| \leq \sqrt{\mathbb{E}(\xi_n - \xi)^2} \rightarrow 0$ . Отже,  $\xi_n \xrightarrow{\mathbb{L}_1} \xi$ .

$$\mathbb{L}_1 \implies \mathbb{P}.$$

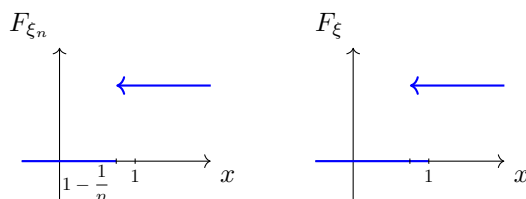
Нехай  $\xi_n \xrightarrow{\mathbb{L}_1} \xi$ . Тоді  $0 \leq \mathbb{P}\{|\xi_n - \xi| > \varepsilon\} \stackrel{\text{н-ть Маркова}}{\leq} \frac{\mathbb{E}|\xi_n - \xi|}{\varepsilon} \rightarrow 0$ , причому  $\forall \varepsilon > 0$ . Отже,  $\xi_n \xrightarrow{\mathbb{P}} \xi$ . ■

**Example 6.2.10** Повернімося до збіжності за розподілом. Точніше відповімо на питання, чому не  $\forall x \in \mathbb{R}$ , а  $\forall x \in \mathbb{R}$ , де  $F_\xi$  – неперервна.

Задамо послідовність  $\xi_n = 1 - \frac{1}{n}$ . Із однієї сторони,  $\xi_n$  – просто сталие число, тобто ми маємо справу з числовою послідовністю. Позначимо  $\xi = 1$ . У нас тоді буде  $\xi_n \rightarrow \xi$  при  $n \rightarrow \infty$ . Ми хочемо зберегти збіжність тепер зі сторони випадкової величини.

З'ясуємо, чому дорівнюють функції розподілу цих випадкових величин:

$$F_{\xi_n}(x) = \begin{cases} 0, & x \leq 1 - \frac{1}{n} \\ 1, & \text{інакше} \end{cases} \quad F_\xi(x) = \begin{cases} 0, & x \leq 1 \\ 1, & \text{інакше} \end{cases}.$$



Зауважимо, що всюди  $F_{\xi_n}(x) \rightarrow F_\xi(x)$ , окрім точки  $x = 1$ . У цьому випадку  $F_{\xi_n}(1) = 1 \not\rightarrow 0 = F_\xi(1)$ . Отже, ми отримали  $\xi_n \xrightarrow{d} \xi$ .

**Remark 6.2.11** Зазначимо, що в збіжностях м.с.,  $\mathbb{L}_2$ ,  $\mathbb{L}_1$ ,  $\mathbb{P}$  фігурувалися  $\xi_n$ ,  $\xi$  в означенні. Але в збіжності  $d$  натомість фігурують саме функції розподілу, самих  $\xi_n, \xi$  нема.

**Remark 6.2.12** Якщо взяти послідовність  $\{\xi_n\}$  – однаково розподілені та незалежні випадкові величини, то інтуїтивно зрозуміло, що в них може не бути м.с.,  $\mathbb{L}_2$ ,  $\mathbb{L}_1$ ,  $\mathbb{P}$  збіжності (оскільки величини незалежні, будемо постійно стрибати зі значенням в інше значення). Але за  $d$  збіжність буде (просто тому що всюди однакові функції розподілу).

**Proposition 6.2.13** Задана послідовність  $\{\xi_n, n \geq 1\}$ . Припустимо, що  $\xi_n \xrightarrow{d} \xi$ , але при цьому  $\xi = \text{const.}$  Тоді  $\xi_n \xrightarrow{\mathbb{P}} \xi$ .

**Proof.**

За умовою,  $F_{\xi_n}(x) \rightarrow F_c(x)$  в точках неперервності, причому функція розподілу  $F_c(x) = \begin{cases} 0, & x \leq c \\ 1, & \text{інакше} \end{cases}$ .

Коротше, збіжності не буде лише при  $x = c$ . Нам треба довести, що  $\mathbb{P}\{|\xi_n - c| \leq \varepsilon\} \rightarrow 1$ .  
 $\mathbb{P}\{\xi_n \in [c - \varepsilon, c + \varepsilon]\} \geq \mathbb{P}\{\xi_n \in [c - \varepsilon, c + \varepsilon]\} = F_{\xi_n}(c + \varepsilon) - F_{\xi_n}(c - \varepsilon) \rightarrow F_c(c + \varepsilon) - F_c(c - \varepsilon) = 1 - 0 = 1$ .  
Отже, дійсно  $\xi_n \xrightarrow{\mathbb{P}} c$ . ■

**Theorem 6.2.14** Критерій середньоквадратичної збіжності до сталої

Задана послідовність  $\{\xi_n, n \geq 1\}$ .

$$\xi_n \xrightarrow{\mathbb{L}_2} c \iff \begin{cases} \mathbb{E}\xi_n \rightarrow c \\ \mathbb{D}\xi_n \rightarrow 0 \end{cases}.$$

**Proof.**

$$\begin{aligned} \xi_n \xrightarrow{\mathbb{L}_2} c &\iff \mathbb{E}(\xi_n - c)^2 \rightarrow 0 \iff \\ &\mathbb{E}(\xi_n - \mathbb{E}\xi_n + \mathbb{E}\xi_n - c)^2 = \mathbb{E}(\xi_n - \mathbb{E}\xi_n)^2 + \mathbb{E}(\mathbb{E}\xi_n - c)^2 + 2\mathbb{E}(\xi_n - \mathbb{E}\xi_n)(\mathbb{E}\xi_n - c) = \mathbb{D}\xi_n + (\mathbb{E}\xi_n - c)^2 + 0 \rightarrow 0 \\ &\iff \begin{cases} \mathbb{D}\xi_n \rightarrow 0 \\ (\mathbb{E}\xi_n - c)^2 \rightarrow 0 \end{cases} \iff \begin{cases} \mathbb{E}\xi_n \rightarrow c \\ \mathbb{D}\xi_n \rightarrow 0 \end{cases}. \end{aligned}$$

**Theorem 6.2.15** Теорема неперервності Леві (окремий випадок)

Задана послідовність  $\{\xi_n, n \geq 1\}$ .

$$\xi_n \xrightarrow{d} \xi \iff \chi_{\xi_n}(t) \rightarrow \chi_{\xi}(t), \forall t \in \mathbb{R}.$$

Без доведення. (TODO:?)

## 6.3 Закон великих чисел

### 6.3.1 Для незалежних та однаково розподілених випадкових величин

**Theorem 6.3.1** Нехай  $\{\xi_n, n \geq 1\}$  – послідовність незалежних та однаково розподілених випадкових величин, причому  $\forall n \geq 1 : \mathbb{E}\xi_n = a$ . Тоді  $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{\mathbb{P}} a$  при  $n \rightarrow \infty$ .

**Proof.**

Для зручності позначу  $\bar{\xi}_n = \frac{\xi_1 + \dots + \xi_n}{n}$ . Запишемо характеристичну функцію для даної випадкової величини:

$$\chi_{\bar{\xi}_n}(t) = \chi_{\frac{\xi_1 + \dots + \xi_n}{n}}(t) = \chi_{\xi_1 + \dots + \xi_n}\left(\frac{t}{n}\right) \stackrel{\text{незалежні}}{=} \chi_{\xi_1}\left(\frac{t}{n}\right) \dots \chi_{\xi_n}\left(\frac{t}{n}\right) \stackrel{\text{однаково розподілені}}{=} \chi^n\left(\frac{t}{n}\right).$$

Розпишемо характеристичну функцію за Тейлором:

$$\chi(h) = \chi(0) + \frac{\chi'(0)}{1!}h + o(h), \quad h \rightarrow 0.$$

$$\chi(h) = 1 + iah + o(h), \quad h \rightarrow 0.$$

$$\chi\left(\frac{t}{n}\right) = 1 + \frac{iat}{n} + o\left(\frac{1}{n}\right), \quad n \rightarrow \infty.$$

$$\chi_{\bar{\xi}_n}(t) = \chi^n\left(\frac{t}{n}\right) = \left(1 + \frac{iat}{n} + o\left(\frac{1}{n}\right)\right)^n, \quad n \rightarrow \infty.$$

$$\ln \chi_{\bar{\xi}_n}(t) = n \ln \left(1 + \frac{iat}{n} + o\left(\frac{1}{n}\right)\right) \sim n \cdot \left(\frac{iat}{n} + o\left(\frac{1}{n}\right)\right) = iat + o(1) \rightarrow iat, \quad n \rightarrow \infty.$$

Отже,  $\chi_{\bar{\xi}_n}(t) \rightarrow e^{iat} = \mathbb{E}e^{iat} = \chi_a(t)$  при  $n \rightarrow \infty$ . Значить, за теоремою Леві,  $\bar{\xi}_n \xrightarrow{d} a$  при  $n \rightarrow \infty$ .

Оскільки гранична випадкова величина стала, то  $\bar{\xi}_n \xrightarrow{\mathbb{P}} a$  при  $n \rightarrow \infty$ . ■

**Theorem 6.3.2 Посилений закон великих чисел Колмогорова**

Нехай  $\{\xi_n, n \geq 1\}$  – послідовність незалежних та однаково розподілених випадкових величин, причому  $\forall n \geq 1 : \mathbb{E}\xi_n = a$ . Тоді  $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{\text{м.ф.}} a$  при  $n \rightarrow \infty$ .

Без доведення.

**6.3.2 Для незалежних та різнорозподілених випадкових величин**

**Theorem 6.3.3** Нехай  $\{\xi_n, n \geq 1\}$  – послідовність незалежних випадкових величин, причому  $\forall n \geq 1 : \mathbb{E}\xi_n = a_n$ . Також нехай  $\forall n \geq 1 : \mathbb{D}\xi_n = \sigma_n^2$ , приому  $\{\mathbb{D}\xi_n, n \geq 1\}$  – рівномірно обмежена. Тоді  $\frac{\xi_1 + \dots + \xi_n}{n} - \frac{a_1 + \dots + a_n}{n} \xrightarrow{\mathbb{L}_2} 0$  при  $n \rightarrow \infty$ .

**Proof.**

Скористаємося критерієм середньоквадратичної збіжності до сталої. Маємо:

$$\mathbb{E} \left( \frac{\xi_1 + \dots + \xi_n}{n} - \frac{a_1 + \dots + a_n}{n} \right) = \mathbb{E} \left( \frac{(\xi_1 - a_1) + \dots + (\xi_n - a_n)}{n} \right) = \frac{\mathbb{E}(\xi_1 - a_1) + \dots + \mathbb{E}(\xi_n - a_n)}{n} = 0 \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Оскільки  $\{\mathbb{D}\xi_n, n \geq 1\}$  – рівномірно обмежені, то  $\exists C > 0 : \forall n \geq 1 : \mathbb{D}\xi_n \leq C$ . Звідси

$$\begin{aligned} \mathbb{D} \left( \frac{\xi_1 + \dots + \xi_n}{n} - \frac{a_1 + \dots + a_n}{n} \right) &= \mathbb{D} \left( \frac{(\xi_1 - a_1) + \dots + (\xi_n - a_n)}{n} \right) \stackrel{\text{незалежність}}{=} \\ &= \frac{\mathbb{D}(\xi_1 - a_1) + \dots + \mathbb{D}(\xi_n - a_n)}{n^2} = \frac{\mathbb{D}\xi_1 + \dots + \mathbb{D}\xi_n}{n^2} = \frac{\sigma_1^2 + \dots + \sigma_n^2}{n^2} \leq \frac{nC}{n^2} = \frac{C}{n} \rightarrow 0 \text{ при } n \rightarrow \infty. \end{aligned}$$

Таким чином, отримали  $\frac{\xi_1 + \dots + \xi_n}{n} - \frac{a_1 + \dots + a_n}{n} \xrightarrow{\mathbb{L}_2} 0$  при  $n \rightarrow \infty$ . ■

**6.4 Центральна гранична теорема****6.4.1 Для незалежних та однаково розподілених випадкових величин**

**Theorem 6.4.1** Нехай  $\{\xi_n, n \geq 1\}$  – послідовність незалежних та однаково розподілених випадкових величин, причому  $\forall n \geq 1 : \mathbb{E}\xi_n = a$  та  $\forall n \geq 1 : \mathbb{D}\xi_n = \sigma^2$ . Тоді  $\frac{(\xi_1 - a) + \dots + (\xi_n - a)}{\sigma\sqrt{n}} \xrightarrow{d} \gamma \in \mathcal{N}(0, 1)$  при  $n \rightarrow \infty$ .

## По-строгому про всі ці поняття

### Дискретна випадкова величина

Визначимо тепер строго, як ми це отримали.

Розглянемо  $P_\xi$  – дискретний ймовірнісний розподіл, тобто існує не більше ніж зліченна підмножина  $S = \{x_1, x_2, \dots\} \subset \mathbb{R}$ , для якої  $P_\xi(X \setminus S) = 0$ , а всі решта  $P_\xi(\{x\}) = p_i$  визначені. Зауважимо, що з іншого боку  $p_i = \mathbb{P}\{\xi = x_i\}$ . Будемо вимагати, щоб події  $\{\xi = x_i\}$  утворювали повну групу подій, тобто  $\sum_{k=1}^{\infty} p_i = 1$ . У такому випадку випадкова величина  $\xi$  буде називатися **дискретною**. Її ймовірнісний розподіл зазвичай записують в той самий вигляд ряд розподілу, який у нас був.

### Неперервна випадкова величина

Визначимо тепер строго, як ми це отримали.

Розглянемо  $P_\xi$  – неперервний ймовірнісний розподіл, тобто  $P_\xi(\{x\}) = 0$  для всіх  $x \in \mathbb{R}$ . Відповідно, звідси отримаємо, що  $\mathbb{P}\{\xi = \{x\}\} = 0$ . Звідси, як ми вже знаємо, функція розподілу  $F_\xi$  буде неперервною на  $\mathbb{R}$ . У такому випадку випадкова величина  $\xi$  буде називатися **неперервною**.

### Абсолютно неперервна випадкова величина та щільність розподілу

Нам неперервної випадкової величини буде недостатньо. Оскільки в кожній одноточковій множині ймовірність нуль, то тоді це виконується й для не більш ніж зліченної множини.

Проте існує множина Кантора. Це борельова множина лебегової міри нуль, яка при цьому більш, ніж зліченна. Тобто ми, наспраді, хочемо захопити більш широкий клас множин, що приведе до звуженню класів неперервних випадкових величин.

Розглянемо  $P_\xi$  – абсолютно неперервний ймовірнісний розподіл відносно  $\lambda_1$ . Тобто тепер  $\forall C$  – борельової множини лебегової міри нуль:  $P_\xi(C) = 0$ . У такому випадку випадкова величина  $\xi$  буде називатися **абсолютно неперервною**.

Ми маємо  $P_\xi \ll \lambda_1$ . Згідно з теоремою Радона-Нікодими, існує функція  $f_\xi: \mathbb{R} \rightarrow \mathbb{R}$ , для якої виконується наступне:

$$P_\xi(B) = \int_B f_\xi(t) d\lambda_1(t), \quad B \in \mathcal{B}(\mathbb{R}).$$

Ось саме ця шукана функція  $f_\xi$  – та сама **щільність розподілу** абсолютно неперервної випадкової величини. Така функція визначена з точністю до еквівалентності  $(\text{mod } \lambda)$ . Такі функції будемо ототожнювати.

Із щойно отриманої формули ми отримаємо властивості щільності розподілу. Покладемо  $B = (-\infty, x]$ .

$$F_\xi(x) = \mathbb{P}\{\xi \leq x\} = \int_{-\infty}^x f_\xi(t) dx - \text{звичне означення щільності.}$$

Якщо ще покласти  $B = \mathbb{R}$ , то отримаємо таку формулу:

$$\int_{-\infty}^{+\infty} f_\xi(t) dt = \mathbb{P}\{\xi \in \mathbb{R}\} = 1.$$

Невід’ємність щільності – наслідок невід’ємності міри.

### Математичне сподівання

**Definition 6.4.2** Математичним сподіванням випадкової величини  $\xi$ , що задана на ймовірнісному просторі  $(\Omega, \mathcal{F}, \mathbb{P})$ , називають такий інтеграл Лебега:

$$\mathbb{E}\xi = \int_{\Omega} \xi(\omega) d\mathbb{P}$$

Математичне сподівання існує, якщо інтеграл збіжний. При існуванні математичного сподівання випадкову величину  $\xi$  ще називають **інтегрованою**.

Зараз будуть такі самі та ще додаткові властивості, які не були зазначені вище.

### Proposition 6.4.3 Властивості математичного сподівання

Справедливе наступне:

1)  $\mathbb{E}|\xi| = 0 \iff \xi = 0$  майже скрізь;

- 2)  $\xi \geq 0$  майже скрізь  $\implies \mathbb{E}\xi \geq 0$ ;
- 3)  $\xi_1 \leq \xi_2$  майже скрізь  $\implies \mathbb{E}\xi_1 \leq \mathbb{E}\xi_2$ ;
- 4)  $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$ ;
- 5)  $\mathbb{E}c = c$  (знову ж таки, стала випадкова величина);
- 6)  $\mathbb{E}\mathbb{1}_A = \mathbb{P}(A)$ ;
- 7)  $\mathbb{E}(\alpha\xi_1 + \beta\xi_2) = \alpha\mathbb{E}\xi_1 + \beta\mathbb{E}\xi_2$ .

**Theorem 6.4.4** Нехай  $\xi$  – випадкова величина на  $(\Omega, \mathcal{F}, \mathbb{P})$  та функція  $g: \mathbb{R} \rightarrow \mathbb{R}$  – борельова. Тоді  $\mathbb{E}g(\xi) = \int_{\mathbb{R}} g(x) dP_{\xi}$ .

**Proof.**

Насправді, рівність  $\int_{\mathbb{R}} g(x) dP_{\xi} = \int_{\Omega} g(\xi(\omega)) d\mathbb{P}$  випливає з теореми про заміну змінної інтеграла Лебега. ■

**Corollary 6.4.5** Якщо  $\xi$  – дискретна випадкова величина та  $g: \mathbb{R} \rightarrow \mathbb{R}$  – борельова, то  $\mathbb{E}g(\xi) = \sum_i g(x_i)p_i$ .

Зокрема звідси отримаємо відому формулу  $\mathbb{E}\xi = \sum_i x_i p_i$ .

**Proof.**

$$\begin{aligned} \mathbb{E}g(\xi) &= \int_{\mathbb{R}} g(x) dP_{\xi} = \int_{\mathbb{R} \setminus S} g(x) dP_{\xi} + \int_S g(x) dP_{\xi} = \int_S g(x) dP_{\xi} = \sum_i \int_{\{x_i\}} g(x_i) dP_{\xi} = \\ &= \sum_i g(x_i)P_{\xi}(\{x_i\}) = \sum_i g(x_i)p_i. \end{aligned}$$

**Corollary 6.4.6** Якщо  $\xi$  – абсолютно неперервна випадкова величина та  $g: \mathbb{R} \rightarrow \mathbb{R}$  – борельова, то  $\mathbb{E}g(\xi) = \int_{\mathbb{R}} g(x)f_{\xi}(x) d\lambda_1(x)$ .

Зокрема звідси отримаємо відому формулу  $\mathbb{E}\xi = \int_{\mathbb{R}} x f_{\xi}(x) dx$ .

**Proof.**

Нам треба довести рівність інтегралів  $\int_{\mathbb{R}} g(x) dP_{\xi} = \int_{\mathbb{R}} g(x)f_{\xi}(x) d\lambda_1(x)$ . Але, насправді, ця рівність працює, оскільки в нас  $P_{\xi} \ll \lambda_1$  та функція  $f_{\xi} = \frac{dP_{\xi}}{d\lambda_1}$ . ■

## Випадкові вектори

Нагадаю, що в нас було трошечки інше означення, яке прив'язувалося до координат.

**Definition 6.4.7** Задано  $(\Omega, \mathcal{F})$  – вимірний простір, також розглянемо  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

Функція  $\xi: \Omega \rightarrow \mathbb{R}^n$  називається **випадковим вектором**, якщо

$$\xi_i, i = \overline{1, n} \text{ – випадкова величина.}$$

Тепер перепишемо еквівалентним чином означення, уже без прив'язки до координат. Це нам дасть підстави узагальнити означення чогось випадкового (величина, вектор або будь-який елемент).

**Definition 6.4.8** Задано  $(\Omega, \mathcal{F})$  – вимірний простір, також розглянемо  $(\mathbb{R}, \mathcal{B}(\mathbb{R}^d))$ .

Функція  $\vec{\xi}: \Omega \rightarrow \mathbb{R}^d$  називається **випадковим вектором**, якщо

$$\xi \text{ – } (\mathcal{F}, \mathcal{B}(\mathbb{R}^d))\text{-вимірна функція.}$$

**Proposition 6.4.9** Вищезгадані два означення справді еквівалентні.

**Proof.**

$\Rightarrow$  Дано: перше означення. Хочемо довести, що  $\vec{\xi} \in (\mathcal{F}, \mathcal{B}(\mathbb{R}^d))$ -вимірною. В курсі теорії міри ми вже якось доводили, що якщо всі  $\xi_i \in \mathcal{F}$ -вимірною, то тоді  $\vec{\xi} \in (\mathcal{F}, \mathcal{B}(\mathbb{R}^d))$ -вимірною (коли мова йшла про композицію відображень).

$\boxed{\Leftarrow}$  Дано: друге означення. Хочемо довести, що кожна  $\xi_i$  – випадкова величина. Тобто хочемо, щоб  $\xi_i$  була  $\mathcal{F}$ -вимірною. Ми розглянемо функцію  $\pi_i$  – проєкція, яка працює як  $\pi_i(\vec{\xi}) = \xi_i$ . Зауважимо, що це неперервне відображення, а тому борельова. Тому  $\xi_i$  буде вимірною як композиція. ■

**Definition 6.4.10** Задано  $(\Omega, \mathcal{F})$  – вимірний простір, також розглянемо вимірний простір  $(U, \mathcal{U})$ . Функція  $\xi: \Omega \rightarrow U$  називається **випадковим елементом**, якщо

$\xi$  –  $(\mathcal{F}, \mathcal{U})$ -вимірна функція.