

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО»

НАВЧАЛЬНИЙ ПОСІБНИК
З ДИСЦИПЛІНИ
“ТЕОРІЯ ЙМОВІРНОСТЕЙ, ЙМОВІРНІСНІ
ПРОЦЕСИ ТА МАТЕМАТИЧНА СТАТИСТИКА”
КУРС ЛЕКЦІЙ
Частина 1

для студентів
спеціальності 126 «Інформаційні системи та технології»

Київ-2021

Навчальний посібник з дисципліни «Теорія ймовірностей, імовірнісні процеси та математична статистика». Курс лекцій. Частина 1 для студентів спеціальності 126 «Інформаційні системи та технології» / Павлов О.А., Гавриленко О.В., Рибачук Л.В. – Київ: КПІ, 2021. – 154 с.

Навчальний посібник призначений для студентів спеціальності 126 «Інформаційні системи та технології» кафедри автоматизованих систем обробки інформації і управління всіх форм навчання. У посібнику наведено систематизоване викладення багатьох основних відомостей з теорії ймовірностей, імовірнісних процесів та математичної статистики з орієнтацією на поглиблене вивчення вказаної дисципліни студентами технічних спеціальностей. Метою авторів було з одного боку закласти фундамент для можливості в майбутньому ознайомлення з сучасною ймовірнісною літературою (аксіоматичний підхід до викладення теорії ймовірностей, строге доведення результатів тощо), а з іншого боку дати більш менш широкий набір результатів для їх застосування в інженерній практиці.

Автори: Павлов Олександр Анатолійович,
доктор технічних наук, професор,
Гавриленко Олена Валеріївна,
кандидат фізико-математичних наук, доцент,
Рибачук Людмила Віталіївна,
кандидат фізико-математичних наук, доцент

Рецензент: Стенін О.А., проф. кафедри ТК ФІОТ, д.т.н

Затверджено
вченою радою
ФІОТ

Протокол № 6
від 20.01.2021 р.

Затверджено
на засіданні кафедри
АСОІУ

Протокол № 9
від 30.12.2020 р.

ЛЕКЦІЯ 1

Основні поняття теорії ймовірностей

Випробування, події

Теорія ймовірностей як наука виникла з глибокої впевненості людства, що за масовими випадковими явищами приховуються певні детерміновані закономірності. Ці закономірності і вивчає теорія ймовірностей.

Наприклад, передбачити наперед, при підкиданні симетричної монети випаде “Герб” чи “Решка”, неможливо. Але, при великій кількості підкидань монети, “Герб” та “Решка” випадають приблизно однакову кількість разів.

Мовою теорії ймовірностей є мова теорії множин. І, як будь-яка математична наука, теорія ймовірностей спирається на первісні означення (аксіоми), які не виводяться, а тільки пояснюються.

Вихідним поняттям теорії ймовірностей є поняття випробування або ймовірнісного (стохастичного) експерименту. **Випробування** – це реалізація певного комплексу умов, яка може повторюватись необмежену кількість разів, в які входять як детерміновані, так і випадкові фактори. Наявність випадкових факторів призводить до неоднозначності результату випробування.

Приклади випробувань: одноразове підкидання симетричної монети або шестигранного грального кубика.

Конкретний результат випробування називають **елементарною подією**.

Приклад. В прикладі з монетою маємо 2 різні елементарні події: “Герб” чи “Решка”, в прикладі з кубиком – 6 елементарних подій: грані з певною цифрою від одного до шести.

Простором елементарних подій, побудованим для даного випробування, (позначають Ω) називається множина всіх різних елементарних подій, кожна з яких може настати внаслідок даного випробування

$$\Omega = \{\omega_1, \dots, \omega_s\}.$$

Приклад. В прикладі з монетою $\Omega = \{Г, Р\}$, в прикладі з кубиком $\Omega = \{\omega_1, \dots, \omega_6\}$.

Складною подією називається будь-яка підмножина простору елементарних подій Ω ($A \subseteq \Omega$). Внаслідок випробування настає складна подія, якщо настала елементарна подія, яка належить (є підмножиною) складній події.

Наслідок. Якщо внаслідок випробування настала конкретна елементарна подія, це означає, що одночасно настали і всі різні складні події, до складу яких належить ця елементарна подія.

Приклад: випробування – одноразове підкидання грального кубика, елементарна подія ω_i – випадання грані з номером i , складна подія $A_1 = \{\omega_1, \omega_3, \omega_5\}$ – випадання грані з непарним номером, складна подія $A_2 = \{\omega_1, \omega_2, \omega_3\}$ – випадання грані, номер якої не перевищує 3.

Серед складних подій відрізняють **достовірну подію** $U = \Omega$ – подію, яка відбудеться завжди внаслідок випробування, тому що вона складається з усіх можливих елементарних подій.

Відрізняють також **неможливу подію** $V = \{\emptyset\}$, яка ніколи не відбудеться внаслідок випробування (не містить жодної елементарної події). За змістом, з нею ототожнюється все, що не може відбутися внаслідок випробування. На мові теорії множин така множина називається порожньою.

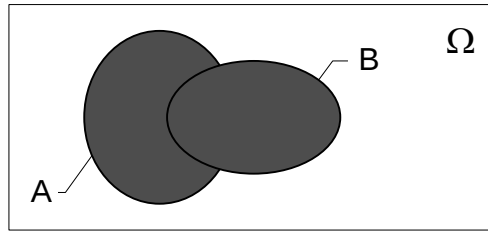
Наприклад: випробування – дворазове підкидання грального кубика, неможлива подія – сумарна кількість очок менше двох.

Для зручності будемо вважати, що $\emptyset \subset \Omega$.

Подія A входить в B ($A \subset B$), якщо всі елементарні події, що входять в A , входять також в B .

Операції над подіями

Подія C є **об'єднанням** A і B ($C = A \cup B$), якщо подія C складається з усіх елементарних подій, що входять в подію A та з усіх елементарних подій, що входять в B . Причому, якщо елементарна подія входить як в A , так і в B , то вона входить в подію C один раз. Тому, що C є підмножиною Ω , то C складається з різних елементарних подій.



$$A \cup B$$

Наслідок. Внаслідок випробування подія C настає тоді і тільки тоді, коли настала подія A або подія B , або та елементарна подія, що входить і в A і в B одночасно.

Вводиться операція об'єднання скінченної чи нескінченно-зліченої кількості подій.

Нехай I – скінченна чи нескінченно-злічена множина індексів, тоді $C = \bigcup_{i \in I} A_i$.

C складається з усіх елементарних подій, що входять в A_i , і якщо елементарна подія входить в різні A_i , більш ніж один раз то в C лише один раз. C настає тільки тоді, коли настала хоча б якась A_i .

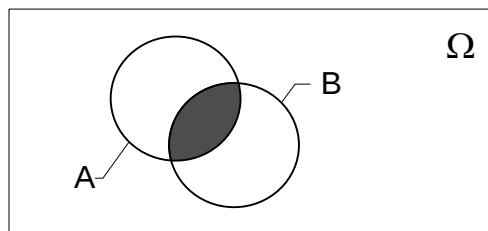
Елементарні тотожності:

$$A \cup A = A$$

$$A \cup \Omega = \Omega$$

$$A \cup V = A$$

Подія C є **перетином** A і B ($C = A \cap B$ або $C = A \cdot B$), якщо подія C складається з усіх елементарних подій, що входять і в A , і в B одночасно.



$$A \cap B$$

Наслідок. Подія C настає тоді і тільки тоді, коли настали і A , і B .

C настає тоді і тільки тоді, коли настали всі події A_i

$$C = \bigcap_{i \in I} A_i.$$

Елементарні тотожності:

$$A \cap A = A$$

$$A \cap \Omega = A$$

$$A \cap V = V$$

Закони:

$$1. A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

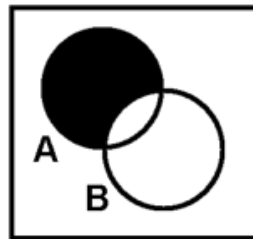
$$2. A \cup (B \cap C) = (A \cup B) \cap C$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$3. A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

ЛЕКЦІЯ 2

Подія $C = A \setminus B$ зветься *різницею подій A і B* якщо вона складається з усіх подій, що входять в A і не входять в B .



Протилежна подія $\bar{A}$ – це така подія, яка відбувається тоді, коли не відбувається вказана подія A : $\bar{A} \cup A = \Omega$, $\bar{A} \cap A = \{\emptyset\}$

A і B несумісні, якщо $A \cap B = \{\emptyset\}$

Наслідок: Якщо A і B несумісні, то вони ніколи не можуть наставати як результат 1 випробування.

Формули де Моргана

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

$$\overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \bar{A}_i \quad \overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \bar{A}_i$$

Складна подія зветься *достовірною* і позначається як $\mathcal{U}$, якщо вона обов'язково настає внаслідок випробування. Протилежною подією до достовірної є неможлива.

$$\mathcal{U} = \Omega \quad V = \bar{\mathcal{U}} = \{\emptyset\}.$$

Частість наставання подій

Нехай простів елементарних подій складається з обмеженої кількості елементарних подій $\Omega = \{\omega_1, \dots, \omega_S\}$.

$$\Omega = \bigcup_{i=1}^S \{\omega_i\}$$

F – позначення множини всіх подій що є всіма можливими різними підмножинами Ω , включаючи неможливу подію $V = \{\emptyset\}$.

Кількість елементів – $|F| = 2^S$.

Беремо довільну подію A , $A \in F$, n разів повторили випробування, для якого побудований простір Ω . n_A – кількість випробувань, в кожному з яких настала подія A .

Частістю наставання події A в n довільних випробувань є $W_n = \frac{n_A}{n}$ (випробування зветься довільним, якщо на нього не накладаються умови).

Частість ще називають *відносною частотою*.

Властивості частоти:

1. Частість належить сегменту від 0 до 1

$$0 \leq W_n(A) \leq 1.$$

2. Частість достовірної події дорівнює одиниці

$$W_n(\mathcal{U}) = 1.$$

3. Якщо $A = \bigcup_{i=1}^K A_i$, $\forall i \neq j, A_i \cap A_j = \{\emptyset\}$ (попарно несумісні), то

$$W_n(A) = \sum_{i=1}^K W_n(A_i)$$

Доведення: провели випробування внаслідок якого настала подія A . За означенням об'єднання це означає, що в цьому випробуванні настала якась подія A_i

($i \in \overline{1; K}$). Кількість випробувань в яких вона настала - n_{A_i} . Оскільки A_i - попарно несумісні, то в цьому випробовуванні не настала жодна інша подія A_j ($j \neq i$). Звідси

$$n_A = \sum_{i=1}^K n_{A_i} \Rightarrow \frac{n_A}{n} = \frac{\sum_{i=1}^K n_{A_i}}{n}.$$

Теоретична ймовірність як формальна наука побудована для випадків, що описують такі випробування, для яких повинно виконуватись:

$$\exists \lim_{n \rightarrow \infty} W_n(A) = \frac{n_A}{n} = P(A)$$

І ця границя однакова для будь-якої серії випробувань і називається **ймовірністю настання події A** .

Здавалося б теорію ймовірності побудувати, взявши за первісне означення $\lim_{n \rightarrow \infty} W_n(A) = P(A)$, але виявилось, що жодна спроба не була прийнята математиками.

Вихід знайшов російський вчений Олексій Миколайович Колмогоров: він ввів аксіоми теорії ймовірності, з яких не випливає, що ймовірність настання події є границею частоті. Цей результат впливає з його теореми: «Посилений закон великих чисел».

Аксіоми теорії ймовірностей Колмогорова (інженерний варіант):

1. $\forall A \subseteq \Omega, 0 \leq P(A) \leq 1$
2. $P(\mathcal{U}) = P(\Omega) = 1$
3. Якщо $A = \bigcup_{i=1}^{K \vee \infty} (A_i), \forall i \neq j, A_i \cap A_j = \{\emptyset\}$, то

$$P(A) = \bigcup_{i=1}^{K \vee \infty} P(A_i).$$

Примітка: Кількість подій може бути обмеженою або нескінченно-зліченою.

Інтуїтивний зміст аксіом Колмогорова: раз від теорії ймовірностей вимагається, що ймовірність настання подій дорівнювала границі частотей, то вони повинні обов'язково задовольняти властивостям частотей.

Ймовірність подій, за Колмогоровим, визначається як кількісна міра, що описує частість настання події у серії довільних випробувань.

Насправді, інженерний варіант аксіом Колмогорова є абсолютно строгим лише для одного конкретного випадку:

$\Omega = \{\omega_1, \dots, \omega_S\}$ – обмежена, або $\Omega = \{\omega_1, \dots, \omega_S \dots\}$ – нескінченна, тобто

$$\Omega = \bigcup_{i=1}^S \omega_i \text{ або } \Omega = \bigcup_{i=1}^{\infty} \omega_i.$$

Існує конструктивний алгоритм, який задає ймовірності по Колмогорову, а з теореми Колмогорова випливає, що цією границею є границя частоти.

$$P(\omega_i) = P_i, 0 \leq P_i \leq 1, \quad i = 1, S \vee \infty$$

$$\sum_{i=1}^{S \vee \infty} P_i = 1, \Omega = \bigcup_{i=1}^{K \vee \infty} \{\omega_i\}, A \subseteq \Omega = \bigcup_{i=1}^{K \vee \infty} \{\omega_{l_i}\}$$

Елементарні події ω_i попарно несумісні за означенням:

$$P(A) = \sum_{i=1}^{K \vee \infty} P_{l_i}.$$

Якщо $A_i \cap A_j = \emptyset, i \neq j$ і $A = \bigcup_{i=1}^K \{A_i\}$, то $P(A) = \sum_{i=1}^K P(A_i)$.

Показати самостійно.

ЛЕКЦІЯ 3

Наслідок з аксіом. Покажемо, що $P(V)=0$.

Дійсно,

$$\Omega \cup \{\emptyset\} = \Omega \quad (V = \{\emptyset\})$$

$$P(\Omega) + P(V) = P(\Omega) \quad (P(\Omega) = 1)$$

Правила порівняння множин:

- Якщо множини містять скінченну кількість елементів, то одна з них більш потужна, якщо кількість її елементів більша, ніж кількість елементів іншої.

Правила порівняння нескінченних множин:

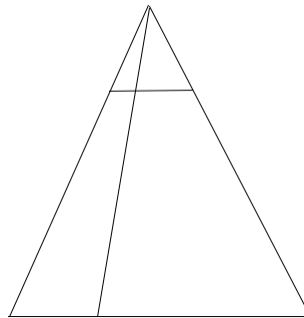
- Дві нескінченні множини мають однакову потужність, якщо між їх елементами можна встановити взаємно однозначну відповідність.
- Одна нескінченна множина більш потужна за іншу, якщо між елементами цих множин не можна встановити взаємно однозначної відповідності, але існує підмножина більш потужної множини, яка має взаємно однозначну відповідність з другою множиною.

Наприклад.

- Перший найменш потужний клас нескінченних множин – це злічені нескінченні множини, ці нескінченні множини еквівалентні натуральному ряду чисел.
- Наступна по потужності множина – множина всіх чисел числової осі чи множина всіх чисел відрізків ненульової довжини.

Дійсно, між множиною всіх чисел числової осі і будь-якою зліченою множиною неможливо встановити взаємно однозначну відповідність. Але, якщо взяти підмножину чисел числової осі, то множина всіх раціональних чисел еквівалентна множині всіх чисел натурального ряду (злічені).

Доводиться, що множина всіх чисел будь-якої числової осі і множина всіх чисел довільного відрізка ненульової довжини еквівалентні.



1. Якщо простір елементарних подій Ω є множина нескінченно-незлічена (більш потужна, ніж нескінченно-злічена множина, наприклад, Ω – числова вісь чи відрізок числової осі), то не існує конструктивного алгоритму побудови $P(A)$, де A – довільна підмножина Ω , і всі $P(A)$ задовольняють аксіоми Колмогорова (інженерний варіант).
2. Існують прилади (один з яких буде розглянуто), коли строго доводиться, що множина подій, кожній з яких ставиться у відповідність ймовірність по Колмогорову і множина цих подій є підмножинами всіх подій Ω , то показано, що $P(A)$, не задовольняє умові задачі.

Таким чином, якщо Ω є нескінченно-незліченою множиною, то множина всіх подій, для кожної з яких існує ймовірність настання по Колмогорову, повинна бути «вужчою» ніж множина всіх підмножин Ω . Для цієї множини задається єдина вимога: будь-яке скінченне чи нескінченне об'єднання чи перетин подій цієї множини повинні належати цій множині (інакше, не існували б граничні теореми теорії ймовірностей).

*****Аксиоми теорії ймовірностей Колмогорова (строгий варіант):**

Множина всіх подій називається **алгеброю** (позначається F), якщо всі її елементи задовольняють умови:

1. $\Omega \in F$
2. $\forall A \in F \Rightarrow \bar{A} \in F$
3. $\forall A \in F, B \in F \Rightarrow A \cup B \in F$

Наслідки.

1. $V \in F$
2. $A \cap B \in F$ (якщо $A, B \in F$)

Доведення:

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

$$A \cap B = \overline{\bar{A} \cup \bar{B}}$$

$$\bigcup_{i=1}^k A_i \in F$$

$$\bigcap_{i=1}^k A_i \in F$$

Довести, використовуючи метод математичної індукції.

3. Якщо A – елемент алгебри, B – елемент алгебри, то $A \setminus B = A \cap \bar{B}$ – елемент алгебри.

Таким чином, алгебра замкнута відносно операцій об'єднання чи перетину довільної обмеженої кількості подій з алгебри.

Приклад. Розглянемо алгебру F півінтервалів: $[a, b)$, $a < b$. В цю алгебру, крім півінтервалів, ще входять всі довільні скінченні об'єднання півінтервалів, що попарно не перетинаються.

В нашому випадку, Ω – це вся числова вісь, тому множину чисел всієї числової осі можна записати як $(-\infty, \infty)$, $[-\infty, \infty)$, $(-\infty, \infty]$, $[-\infty, \infty]$

$$\bigcap_{n=1}^{\infty} [a, a + \frac{1}{n}) = a \notin F$$

$$\bigcap_{n=1}^{\infty} [a - \frac{1}{n}, a) = \{\emptyset\} \in F$$

Множина подій називається **сігма-алгеброю** (σ -алгеброю), якщо:

1. $\Omega \in \sigma$
2. $\forall A \in \sigma \Rightarrow \bar{A} \in \sigma$
3. $\forall A_i, i = \overline{1, \infty} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \sigma$

Наслідки.

- За формулами де Моргана можна довести, що довільний нескінченний перетин є σ -алгеброю, тобто σ -алгебра замкнута відносно довільного нескінченного об'єднання чи перетину.
- σ -алгебра замкнута відносно довільного скінченного об'єднання чи перетину.

ЛЕКЦІЯ 4

***Приклади σ -алгебр

1. Клас підмножин Ω , що складається з двох подій $\{U, V\} = \{\Omega, \emptyset\} \in \sigma$ -алгеброю.
2. Клас всіх підмножин множини $\Omega \in \sigma$ -алгеброю.

3. Борелівська алгебра

Нехай K – деякий непорожній клас підмножин Ω . Тоді σ -алгебра називається *мінімальною* $\sigma(K)$, якщо вона містить K , яка задовольняє умовам:

- $K \subseteq \sigma(K)$
- $\forall \sigma, K \subseteq \sigma \Rightarrow \sigma(K) \subseteq \sigma$

Теорема 1: Якщо K не була множина K , існує єдина мінімальна σ алгебра, що містить K .

Доведення.

Розглянемо перетин всіх σ -алгебр, що містять K . Доведемо, що цей перетин і є мінімальною σ -алгеброю, що містить K .

Перетин існує, якщо є хоч одна σ -алгебра, що містить K (σ -алгебра всіх підмножин множини Ω).

По-перше $K \subseteq \cap \sigma(K)$.

По-друге перетин довільного числа σ -алгебр є σ -алгеброю (показати самим). Таким чином, перетин $\sigma(K)$ є мінімальною σ -алгеброю.

Нехай $\Omega \in (-\infty; +\infty)$ – числова вісь.

$$K = \{ \forall [a, b) \}, \text{ де } a < b$$

Мінімальна σ -алгебра, що містить множину всіх підінтервалів числової осі існує і називається *борелівською алгеброю*.

Вона не співпадає з множиною всіх підмножин дійсної осі, але містить у собі достатньо великий клас множин, з якими ми маємо справу при вирішенні практичних задач.

Теорема 2.

Розглянемо множини K_1, K_2, K_3, K_4

$$K_1 = \{ \forall [a, b) \}, \text{ де } a < b$$

$$K_2 = \{ \forall (a, b] \}, \text{ де } a < b$$

$$K_3 = \{ \forall (a, b) \}, \text{ де } a < b$$

$$K_4 = \{ \forall [a, b] \}, \text{ де } a \leq b$$

$$\sigma(K_1) = \sigma(K_2) = \sigma(K_3) = \sigma(K_4) = \mathcal{B}$$

Таким чином $\mathcal{B}$ містить всі напівінтервали, всі інтервали, сегменти, числа, скінченні та нескінченні об'єднання чи перетини попередніх подій, всі події, що їм протилежні.

Теорема 3 (якісна): Борелівська алгебра є власною підмножиною множини всіх підмножин чисел осі.

Нехай на Ω задана алгебра F , тоді **числова скалярна функція** $P(\cdot)$, аргументом якої є подія, що входить в алгебру F (множина, елементами якої є події, що задовольняє умовам алгебри) зветься **скінчено-адитивною мірою**, якщо вона задовольняє умовам:

- 1) $\forall A \in F, P(A) \geq 0$
- 2) $P(\Omega) = 1$
- 3) Якщо $A \cap B = \{\emptyset\}, A \in F, B \in F$, то $P(A \cup B) = P(A) + P(B)$

Наслідки:

- 1) $P(\emptyset) = 0$
 $U \cup V = U$
 $1 + P(V) = 1$
- 2) $P(A) \leq 1$
 $A \cup \bar{A} = U$
 $P(A) + P(\bar{A}) = 1$
- 3) $P(\bigcup_{i=1}^k A_i) = \sum_{i=1}^k P(A_i)$

Останнє твердження доводиться за допомогою методу математичної індукції.

Є алгебра F задана на Ω , тоді числова скалярна функція $P(\cdot)$ зветься **злічено-адитивною мірою** на F , якщо виконується:

- 1) $\forall A \in F \quad P(A) \geq 0$
- 2) $P(\Omega) = 1$
- 3) Якщо $\bigcup_{i=1}^{\infty} A_i \in F, \forall A_i \cup A_j, i \neq j$
 $\forall A_i \in F$, тоді
 $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Наслідок: злічено-адитивна міра є скінчено-адитивною мірою (довести самим).

ЛЕКЦІЯ 5

Є простір елементарних подій Ω на якому задано σ -алгебру.

Числова скалярна функція, аргументом якої є довільна подія, що є елементом алгебри, називається **зліченно-адитивною ймовірнісною мірою**, що задана на σ -алгебрі, якщо для неї виконуються наступні умови:

$$1) P(A) \geq 0, \quad \forall A \in \sigma$$

$$2) P(\Omega) = 1$$

$$3) \forall A_i, i = \overline{1, \infty}, \quad \forall i \neq j \quad A_i \cap A_j = \{\emptyset\}$$

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

Наслідок.

$$\forall A_i, i = \overline{1, s}, \quad \forall i \neq j \quad A_i \cap A_j = \{\emptyset\}$$

$$P(\bigcup_{i=1}^s A_i) = \sum_{i=1}^s P(A_i).$$

Показати самостійно.

Первісне означення, з якого починається математична теорія ймовірностей – ймовірнісний простір.

Ймовірнісним простором називається трійка (Ω, σ, P) , де Ω – простір елементарних подій, на якому задана множина подій, що є σ -алгеброю, на якій задана зліченно-адитивна міра P .

Теорема Каратеодорі (дається без доведення).

Нехай на Ω задана множина подій, що є алгеброю F . На алгебрі F задана зліченно-адитивна міра $Q(\cdot)$. Тоді на мінімальній σ -алгебрі, що містить алгебру F , існує лише одна зліченно-адитивна ймовірнісна міра $P(\cdot)$, що задовольняє умові $\forall A \in F \quad Q(A) = P(A)$.

Інша назва: **теорема про однозначне продовження міри** $Q(\cdot)$ на мінімальну σ -алгебру, що містить F .

Приклади ймовірнісних просторів.

Приклад 1 (див. приклад після інженерного варіанту аксіом Колмогорова)

$$\Omega = \{w_1, \dots, w_s\} \text{ чи } \Omega = \{w_1, \dots, w_s, \dots\}$$

$$P(w_i) = P_i \geq 0, \quad \sum_{i=1}^{s \vee \infty} P_i = 1$$

$A = \bigcup_{l=1}^{k \vee \infty} w_{i_l}$ (w_{i_l} – елементарна подія, що входить в множину A , яка є підмножиною Ω)

$$P(A) = \sum_{l=1}^{k \vee \infty} P_{i_l} \quad \forall w_{i_l} \in A$$

Показали самим, що для σ , що є множиною всіх підмножин Ω , ми побудували зліченно-адитивну ймовірнісну міру P . Таким чином, ми задали ймовірнісний простір для прикладу 1.

Приклад 2 (Класичне означення ймовірності)

$\Omega = \{w_1, \dots, w_s\}$ – складається з обмеженої кількості елементарних подій, які є *рівноправними*, тобто до випробувань, дослідник не може стверджувати, що якась з цих елементарних подій наставатиме частіше, ніж інша.

$$P(w_i) = P_i = \text{const} \quad \forall i$$

$$P_i = \frac{1}{s}$$

$$P(A) = \frac{k}{s}$$

Якщо елементарні події рівноймовірні, то ймовірність довільної події дорівнює дробу: *чисельник* – кількість елементарних подій, що входять в дану подію, *знаменник* – загальна кількість всіх елементарних подій.

***Приклад 3 (кидання голки навмання в півінтервал $[0, 1)$)

Ω – множина всіх чисел півінтервала $[0, 1)$.

$[0, 1)$ – щоб далі не проводити процедуру нормування.

Беремо півінтервал, а не сегмент чи інтервал, оскільки в нашому курсі ми ввели борелівську алгебру – мінімальна σ -алгебра, що містить множину всіх півінтервалів.

S - площа вістря голки $= 0$, тобто результат випробування – поява довільного числа з півінтервалу $[0, 1)$.

Множина всіх елементарних подій — це множина всіх чисел з півінтервалу $[0, 1)$.

$$\Omega = [0, 1)$$

З означення випробування можна вважати, що всі елементарні події рівноймовірні. Значить ймовірність настання будь-якого числа:

$$P(x) = \text{const} \quad \forall x \in [0, 1)$$

$$1) P(x) > 0$$

$$2) P(x) = 0$$

Нехай вірно 1).

$$P(x) = C > 0$$

$$A \subseteq B \quad P(A) \leq P(B)$$

$$B = A \cup (B \setminus A)$$

$$P(B) = P(A) + P(B \setminus A) \quad (\text{кожен доданок} - \text{невід'ємний})$$

Нехай B – множина всіх чисел півінтервалу $[0, 1)$, A – всі раціональні числа півінтервалу $[0, 1)$, їх нескінченно-злічена кількість, тобто ймовірності їх об'єднання можна просумувати.

$$\text{Нехай має місце перше твердження: } P(x) = C > 0$$

тоді $P(A) = \infty$, що суперечить 2-ій аксіомі. $P(x)$ не може дорівнювати додатньому числу. Таким чином, $\forall x \quad P(x) \equiv 0$.

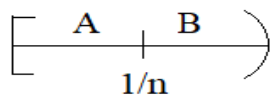
Примітка. Далі покажемо, що для випадку коли Ω – нескінченно-незліченні множини, то події, які можуть настати і мають ймовірність нуль чи події, які можуть не настати і мають ймовірність одиницю в інженерній теорії ймовірностей трактуються наступним чином:

- Якщо подія може настати і має ймовірність 0, то вона ніколи не настає в будь-якій обмеженій серії випробування.
- Якщо подія може не настати і має ймовірність 1, то ця подія завжди настає в будь-якій скінченній кількості випробувань.

Складна подія, що складалася з обмеженої чи нескінченно-зліченої кількості чисел, має ймовірність настання 0.

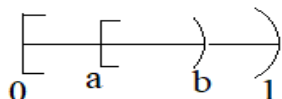
Виникає питання: які ж події мають ненульову ймовірність?

Розглянемо половину півінтервалу.



$$P(A) = P(B) = \frac{1}{2} \text{ (імовірність потрапляння в 1-у чи 2-у}$$

половини півінтервалу).



Розглянемо будь-який півінтервал.

Імовірність потрапляння дорівнює $b - a$.

Таким чином, ненульова ймовірність настання може бути лише в складних подіях, що складаються з нескінченно-незліченої кількості чисел.

На півінтервалі $[0, 1)$ задамо алгебру півінтервалів F , яка складається з усіх різних півінтервалів, кожен з яких входить в $[0, 1)$, і будь-якого скінченного об'єднання попарно-несумісних півінтервалів, що входить в $[0, 1)$ (попарно несумісних).

Задамо на алгебрі півінтервалів F скінченно-адитивну міру

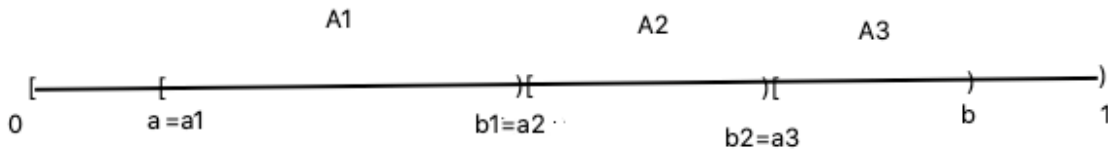
$$P([a; b)) = b - a, \quad [a; b) \in [0; 1)$$

$$P\left(\bigcup_{i=1}^k [a_i; b_i)\right) = \sum_{i=1}^k b_i - a_i, \quad \forall i \neq j \quad [a_i; b_i) \cap [a_j; b_j) = \emptyset$$

ЛЕКЦІЯ 6

Покажемо, що ця скінченно-адитивна міра є зліченою мірою на алгебрі напівінтервалів на інтервалі $[0, 1)$.

Строгу теорему не доводимо, наводимо обґрунтування ідеї доведення.



$$\bigcup_{i=1}^{\infty} A_i = [a, b] \in \mathcal{F}.$$

Ймовірність попадання в інтервал:

$$P(\bigcup_{i=1}^{\infty} A_i) = b - a = \sum_{i=1}^{\infty} (b_i - a_i),$$

Дійсно, обмежена, монотонно зростаюча числова послідовність має границю.

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n (b_i - a_i) = \sum_{i=1}^{\infty} (b_i - a_i) = b - a \quad (\text{при } n \rightarrow \infty)$$

$$P([a, b]) = b - a.$$

Таким чином довели, що наша міра є злічено-адитивною на інтервалі $[0, 1]$.

Знаходимо мінімальну σ -алгебру на інтервалі $[0, 1]$, що містить $\mathcal{F}$.

За доведеною раніше теоремою, мінімальна σ -алгебра існує і називається **борелівською алгеброю на напівінтервалі $[0, 1]$** .

$$\sigma(\mathcal{F}) = \mathcal{B}_{[0, 1]}$$

$$\{[0, 1], \mathcal{B}_{[0, 1]}, P\}$$

Таким чином, за теоремою Каратеодорі побудували злічено-адитивну міру, що побудована на Борелівській алгебрі $\mathcal{B}$, яка містить алгебру напівінтервалів.

Для всіх подій, що входять у Борелівську алгебру і в алгебру інтервалів, їх ймовірності приймають однакове значення.

Теорема (без доведення).

Якщо взяти множину всіх підмножин чисел напівінтервалу $[0, 1]$ замість $\mathcal{B}$, то не існує жодної злічено адитивної ймовірнісної міри, заданої на множині всіх підмножин чисел інтервалу $[0, 1]$, для якої б виконувалось:

$$\forall a < b, P([a; b]) = b - a.$$

Приклад.

$$\Omega = (x_i, i = \overline{1, \infty})$$

Всі елементарні події рівноймовірні.

Ця задача не розв'язується в рамках теорії імовірності, оскільки не виконуються аксіоми Колмогорова.

$P(x_i)$ може дорівнювати нулю, або якомусь заданому числу.

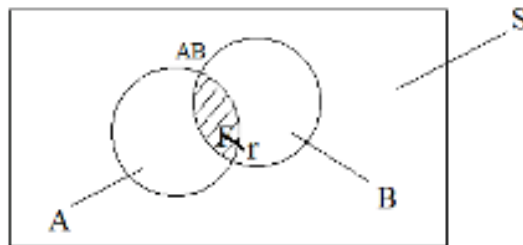
В першому випадку $P(\Omega)=0$, в другому $P(\Omega)=\infty$ (не виконується друга аксіома).

Умовна імовірність $P(A/B)$

Проводиться деяке випробування. Відомо, що внаслідок цього випробування настала подія B (елементарна подія, що належить $\mathbf{B}$). Яка ймовірність того, що внаслідок даного випробування настане подія A ?

Розглянемо приклад:

$$\Omega = \{\omega_1, \dots, \omega_s\}, P(\omega_i) = \frac{1}{s}, i = \overline{(1, s)}.$$



Тоді $P(A/B) = \frac{k}{r} = \frac{k/s}{r/s} = P(A \cdot B)/P(B)$, де k – кількість елементарних подій, що входять в $(A \cdot B)$.

Дійсно, якщо в випробуванні настала подія B , то це означає, що для цього випробування Ω перетворилася в B , отже, кількість всіх елементарних подій – r , а ймовірність настання кожної з них дорівнює $\frac{1}{r}$.

Подія A перетворилась в $(A \cdot B)$, що складається з k елементарних подій.

$$P(A/B) = k/r = P(A \cdot B)/P(B)$$

Це єдиний випадок, для якого виводиться формула умовної ймовірності.

Обґрунтування цієї формули для загального випадку

Проведемо n випробувань. Нехай n_B – кількість випробувань, в кожному з яких настала подія B . А в n_{AB} випадках настала A і B одночасно. Тоді, за означенням умовна частість настання події A , за умови, що настала подія B , дорівнює

$$W_n(A/B) = n_{AB}/n_B.$$

Розділивши чисельник і знаменник на n , отримуємо

$$\frac{n_{AB}/n}{n_B/n} = \frac{Wn(AB)}{Wn(B)}$$

Оскільки далі буде показано, що частість настання події дорівнює ймовірності настання події за Колмогоровим для необмежено великого n , то формула для відносних частот переходить до формули відносних ймовірностей.

Якщо $P(A/B) = \frac{P(AB)}{P(B)}$, $P(B/A) = \frac{P(AB)}{P(A)}$, то

$$P(A \cdot B) = P(A) \cdot P(B/A)$$

$$P(B \cdot A) = P(B) \cdot P(A/B)$$

Покажемо, що

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_k) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cdot A_2) \cdot \dots \cdot P(A_k/(A_1 \cdot A_2 \cdot \dots \cdot A_{k-1})) \quad (6.1),$$

якщо

$$P(\cap_{i=1}^{k-1} A_i) \neq 0. \quad (6.2)$$

Доведення.

Оскільки справедливі включення:

$$\cap_{i=1}^{k-1} A_i \subseteq \cap_{i=1}^{k-2} A_i \subseteq \dots \subseteq A_1$$

Тоді з (6.2) випливає, що формула, яку ми хочемо довести, може існувати.

В цьому випадку в кожній умовній ймовірності знаменник не дорівнює нулю.

Далі використовуємо метод математичної індукції:

1. Для двох подій справедлива рівність $P(A \cdot B) = P(A) \cdot P(B/A)$ – доведено вище.
2. Припускаємо, що (6.1) справедливо для k подій.
3. Доведемо справедливість рівності для $k+1$ подій.

Позначимо $A_1 \cdot A_2 \cdot \dots \cdot A_k = A$, $A_{k+1} = B$, тоді

$$P(A \cdot B) = P(A) \cdot P(B/A).$$

$$\begin{aligned} P(A_1 \cdot A_2 \cdot \dots \cdot A_k) \\ = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cdot A_2) \cdot \dots \cdot P(A_k/(A_1 \cdot A_2 \cdot \dots \cdot A_{k-1})). \end{aligned}$$

Тобто (6.1) доведено.

ЛЕКЦІЯ 7

Виведемо формулу:

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

Доведення.

$$\Omega = (A \cup B) \cup (\overline{A \cup B})$$

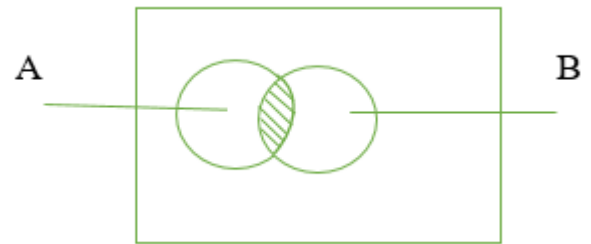
$$P(A \cup B) = 1 - P(\overline{A \cup B})$$

$$\Omega = A \cup (\overline{A} \cdot (B \cup \overline{B}))$$

$$1 - P(\overline{A} \cdot \overline{B}) = P(A) + P(\overline{A} \cdot B)$$

$$B = B \cdot (A \cup \overline{A})$$

$$P(\overline{A} \cdot B) = P(B) - P(A \cdot B)$$



Незалежні події

Дві події A і B називаються **незалежними**, якщо ймовірність їх перетину дорівнює добутку їх ймовірностей

$$P(A \cdot B) = P(A) \cdot P(B).$$

Якщо ця рівність не виконується, то події називаються **залежними**.

Наслідки. Якщо $P(A) \neq 0$, $P(B) \neq 0$, то маємо:

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

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

Нехай $P(A/B) = P(A)$ або $P(B/A) = P(B)$ ($P(A) \neq 0$, $P(B) \neq 0$), тоді події *незалежні*.

Дійсно, якщо A і B незалежні та $P(B) > 0$, то

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A) \Rightarrow P(A \cap B) = P(A) \cdot P(B).$$

$$\text{Якщо ж відомо, що } P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A),$$

то $P(A \cap B) = P(A)P(B)$, а це означає, що A і B незалежні.

Таким чином, для подій ненульової ймовірнісної міри маємо еквівалентне означення незалежності: дві події A і B незалежні, коли умовна ймовірність настання однієї з них, за умови, що друга відбулась, дорівнює безумовній ймовірності першої події.

Теорема: Якщо A і B незалежні, то A і $\bar{B}$, $\bar{A}$ і B , $\bar{A}$ і $\bar{B}$ також незалежні.

Доведення.

Доведемо, що A і $\bar{B}$ незалежні (решту довести самостійно). Зауважимо, що

$$A \bar{B} \cup AB = A, \quad P(A \bar{B}) + P(AB) = P(A).$$

Тому $P(A \bar{B}) = P(A) - P(AB)$.

Якщо A і B незалежні, то

$$P(A \bar{B}) = P(A) - P(A)P(B) = P(A)[1 - P(B)] = P(A) P(\bar{B}).$$

Отже, A і $\bar{B}$ незалежні.

Довести самостійно, що у випадку, коли A і B незалежні та $P(A) > 0$, $P(B) > 0$, то A і B обов'язково сумісні.

Маємо k подій $A_1, A_2, \dots, A_k$. Події $A_1, A_2, \dots, A_k$ незалежні, якщо

$$P\left(\bigcap_{i=1}^k A_i\right) = \prod_{i=1}^k P(A_i).$$

Означення. Випадкові події $A_1, A_2, \dots, A_k$ — **незалежні в сукупності**, якщо для них виконується $\forall r \leq k$ та для довільного набору індексів $i_1, i_2, \dots, i_r$

$$P\left(\bigcap_{l=1}^r A_{i_l}\right) = \prod_{l=1}^r P(A_{i_l})$$

Зокрема, якщо $A_1, A_2, \dots, A_k$ незалежні у сукупності, то довільні дві події A_i і A_j ($i \neq j$) незалежні.

Наслідок. З незалежності подій $A_1, A_2, \dots, A_k$ не випливає їх незалежність в сукупності.

Нехай події $A_1, A_2, \dots, A_k$ незалежні в сукупності

$$\forall r \leq k \quad j \notin \{i_1, i_2, \dots, i_r\}$$

$$P(A_j / A_{i_1} A_{i_2} \dots A_{i_r}) = P(A_j). \quad (7.1)$$

Дійсно, з незалежності у сукупності $\{A_i\}$ маємо:

$$P(A_j / A_{i_1} A_{i_2} \dots A_{i_k}) = \frac{P(A_j A_{i_1} A_{i_2} \dots A_{i_k})}{P(A_{i_1} A_{i_2} \dots A_{i_k})} = \frac{P(A_j) \prod_{r=1}^k P(A_{i_r})}{\prod_{r=1}^k P(A_{i_r})} = P(A_j).$$

Навпаки, якщо для довільних подій з $\{A_i\}$ має місце (1), то $A_i, i=\overline{1, n}$, незалежні у сукупності (довести самостійно, використавши формулу (6.1) для будь-якої підмножини подій з множини подій $A_1, A_2, \dots, A_k$).

Формула повної ймовірності

Розглянемо систему k подій $B_1, B_2, \dots, B_k$, які ще називаються *гіпотезами*.

Нехай $B_i \cap B_j = \emptyset, \forall i \neq j$, тобто події системи попарно несумісні, та нехай $P(B_i) > 0, i=\overline{1, k}$.

Нехай для події A має місце рівність $A \cap (\bigcup_{i=1}^k B_i) = A$ (об'єднання B_i накриває A).

Якщо у випробуванні настала подія A , то у цьому випробуванні обов'язково настала одна з подій $B_1, \dots, B_k$.

Тоді має місце **формула повної ймовірності**

$$P(A) = \sum_{i=1}^k P(B_i) P(A / B_i).$$

Дійсно, якщо $B_1, B_2, \dots, B_k$ несумісні, то несумісні $A \cdot B_1, A \cdot B_2, \dots, A \cdot B_k$. Тоді з третьої аксіоми теорії ймовірностей та формули множення ймовірностей виходить

$$P(A) = P(A \bigcup_{i=1}^k B_i) = P(\bigcup_{i=1}^k A B_i) = \sum_{i=1}^k P(A B_i) = \sum_{i=1}^k P(B_i) P(A / B_i).$$

Формула повної ймовірності використовується для розв'язку задач, де формулу класичної ймовірності не можна використати.

Приклад.

Маємо урни трьох складів.

- 1) 5 урн складаються з 6 білих та 5 чорних куль.
- 2) 3 урни складаються з 10 білих та 1 чорної кулі.
- 3) 7 урн складаються з 1 білої та 10 чорних куль.

Навмання вибирається урна, а потім навмання з неї виймається куля.

Визначити ймовірність, що ця куля біла.

Вводимо гіпотези (вони несумісні і їх сума дорівнює Ω):

B_1 – вибрана куля з довільної урни першого складу;

B_2 – вибрана куля з довільної урни другого складу;

B_3 – вибрана куля з довільної урни третього складу.

Подія A – всі білі кулі. Обчислюємо ймовірність гіпотез та ймовірність події A за умови, що мала місце та чи інша гіпотеза:

$$P(B_1) = \frac{5}{15}; \quad P(B_2) = \frac{3}{15}; \quad P(B_3) = \frac{7}{15};$$

$$P(A/B_1) = \frac{6}{11}; \quad P(A/B_2) = \frac{10}{11}; \quad P(A/B_3) = \frac{1}{11}.$$

За формулою повної ймовірності:

$$P(A) = P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + P(B_3)P(A/B_3) = \frac{5}{15} \cdot \frac{6}{11} + \frac{3}{15} \cdot \frac{10}{11} + \frac{7}{15} \cdot \frac{1}{11} = \frac{67}{165}.$$

Формула Байєса

Постановка задачі схожа на ту, що приводила до використання формули повної ймовірності, але є суттєві відмінності в постановці запитання. Докладніше, знову є система k подій $B_1, B_2, \dots, B_k$. Події несумісні, тобто $B_i \cap B_j = \emptyset, i \neq j$, та мають ненульові ймовірності $P(B_i) > 0, i = \overline{1, k}$. Є подія A з ненульовою ймовірністю $P(A) > 0$, яка задовольняє умові $A \cap \left(\bigcup_{i=1}^k B_i \right) = A$. Відомо, що в результаті випробування подія A реалізувалась. Треба обчислити умовні ймовірності $P(B_i/A), i = \overline{1, k}$. Ймовірність подій $P(B_i), i = \overline{1, k}$, називаються апріорними (до випробування). Умовні ймовірності $P(B_i/A)$,

$i=\overline{1,k}$, називаються апостеріорними (після випробування) ймовірностями подій B_i , $i=\overline{1,k}$.

Відповідь на постановку задачі дає **формула Байєса**

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{\sum_{j=1}^k P(B_j)P(A/B_j)}, \quad i=\overline{1,k}.$$

Доведення просте. За формулою множення ймовірностей:

$$P(AB_i) = P(A) \cdot P(B_i/A) = P(B_i) \cdot P(A/B_i).$$

Звідси

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{P(A)},$$

а далі в знаменнику треба використати формулу повної ймовірності і одержати формулу Байєса.

Приклад: Умови задачі ті ж, що і в попередньому прикладі. Але тепер відомо, що після вибору урни і витягування з неї кулі вона виявилась білою (перша відмінність), тобто подія A реалізувалась в результаті випробування. Треба обчислити ймовірності $P(B_i/A)$, $i=\overline{1,3}$, того, що була вибрана урна першого, другого, третього складу (друга відмінність).

Застосуємо формулу Байєса, використовуючи вже підраховані ймовірності. Одержуємо:

$$P(B_1/A) = \frac{P(B_1)P(A/B_1)}{P(A)} = \frac{\frac{5}{15} \cdot \frac{6}{11}}{\frac{67}{165}} = \frac{30}{67};$$

$$P(B_2/A) = \frac{P(B_2)P(A/B_2)}{P(A)} = \frac{\frac{3}{15} \cdot \frac{10}{11}}{\frac{67}{165}} = \frac{30}{67};$$

$$P(B_3/A) = \frac{P(B_3)P(A/B_3)}{P(A)} = \frac{\frac{7}{15} \cdot \frac{1}{11}}{\frac{67}{165}} = \frac{7}{67}.$$

ЛЕКЦІЯ 8

Композиція двох випробувань

Маємо два випробування:

$$1. \Omega_1 = \{E_1 \dots E_{m_1}\}$$

$$P_1(E_i) \quad i = \overline{1, m_1}$$

$$\sum_{i=1}^{m_1} P_1(E_i) = 1$$

$$(\Omega_1 \sigma_1 P_1)$$

$$2. \Omega_2 = \{Q_1 \dots Q_{m_2}\}$$

$$P_2(Q_j) \quad j = \overline{1, m_2}$$

$$\sum_{j=1}^{m_2} P_2(Q_j) = 1$$

$$(\Omega_2 \sigma_2 P_2)$$

Композицією двох випробувань зветься проведення першого і другого випробування.

Чи одночасно проводяться випробування, чи між ними є проміжок часу – не має значення. Це задається конкретною задачею.

Таке означення не суперечить загальному означенню випробування.

Побудуємо ймовірнісний простір для композиції двох випробувань.

Розглянемо декартів добуток $A \oplus B$, де $E_i \in A$, $Q_j \in B$. За його означенням, до $A \oplus B$ входять всі пари (E_i, Q_j) , такі, що $E_i \in A$ і $Q_j \in B$.

$$\Omega = \Omega_1 \oplus \Omega_2 = \{E_i Q_j\}, \text{ де } i = \overline{1, m_1}, j = \overline{1, m_2}.$$

Всіх можливих пар $m_1 \cdot m_2$ (обмежена кількість), тоді σ -множина усіх підмножин з Ω , де P задається числами $P(E_i Q_j) \forall i = \overline{1, m_1}, j = \overline{1, m_2}$.

В загальному випадку, $P(E_i Q_j)$ не задається ймовірностями $P_1(E_i), P_2(Q_j)$.

1) Приклад:

1. Випробування - підкидання шестигранного кубика. Елементарна подія Ω_1 – випадання конкретного числа $E_i = \{i, i = \overline{1,6}\}$.

$$P(E_i) = \frac{1}{6}.$$

2. Друге випробування є першим випробуванням, результат якого є квадрат числа, що настало в першому випробуванні.

$$P(Q_j) = \frac{1}{6}, Q_j = \{j^2, j = \overline{1,6}\}.$$

Робимо композицію першого і другого випробувань:

$$\Omega = \Omega_1 \oplus \Omega_2,$$

$$P(ii^2, i = \overline{1,6}) = \frac{1}{6}, P(ij^2, i \neq j) = 0.$$

Приклад: Перше і друге випробування – ті самі, що в попередньому прикладі, тільки кубики кидають різні люди в різних аудиторіях.

Ω_1, Ω_2 - ті самі,

Робимо композицію першого і другого випробувань:

$$P(ij^2, \forall i \neq j) = \frac{1}{36}.$$

Дійсно, по $P_1(E_i), P_2(Q_j)$ не можна знайти $P(E_i Q_j)$, тому що ми не знаємо, які випадкові фактори входять в 1 випробування, і які входять у 2 випробування.

В 1-му прикладі випадкові фактори обох випробувань однакові, а в 2-му прикладі випробування не мають спільних випадкових факторів.

Можливий проміжний варіант: є спільні та різні випадкові фактори двох випробувань.

1). Два випробування називаються **незалежні**, якщо

$$\forall i, j P(E_i Q_j) = P_1(E_i) \cdot P_2(Q_j).$$

2). Два випробування називаються **незалежні**, якщо в їх склад входять різні випадкові фактори

Покажемо, що з означення 2 випливає означення 1.

Позначимо $A = E_i \subset \Omega_1$

$$A' = \bigcup_{j=1}^{m_2} E_i Q_j \subset \Omega.$$

З означення випливає, що A' настане в результаті композиційного випробування тоді і тільки тоді, коли в першому випробуванні настане E_i

Вводимо симетрично:

$$B = Q_j \subset \Omega_2$$

$$B' = \bigcup_{j=1}^{m_2} E_i Q_j \subset \Omega$$

B' настає, коли у другому випробуванні настало Q_j

$$A' \cap B' = \{E_i Q_j\} = A \oplus B$$

$$A = \bigcup_{l=1}^{s_1} E_{i_l} \subset \Omega_1$$

$$A' = \bigcup_{l=1}^{s_1} \bigcup_{j=1}^{m_2} \{E_{i_l} Q_j\} = A \oplus \Omega_2.$$

A' настає, тоді і тільки тоді, коли у першому випробуванні настане складні подія A

Симетрично

$$B = \bigcup_{t=1}^{s_2} Q_{j_t} \subset \Omega_2$$

$$B' = \bigcup_{i=1}^{m_1} \bigcup_{t=1}^{s_2} \{E_i Q_{j_t}\} \subset \Omega$$

$$B' = \Omega_1 \oplus B$$

B' настає, тоді і тільки тоді, коли у другому випробуванні настала складні подія B

$$A' \cap B' = A \oplus B = \bigcup_{l=1}^{s_1} \bigcup_{t=1}^{s_2} E_{i_l} Q_{j_t}$$

$$P(A') = P_1(A)$$

$$P(A') = \sum_{l=1}^{s_1} \sum_{j=1}^{m_2} P(E_{i_l} Q_j) = P_1(A) = \sum_{l=1}^{s_1} P_1(E_{i_l})$$

$$P(B') = \sum_{i=1}^{m_1} \sum_{t=1}^{s_2} P(E_i Q_{j_t}) = P_2(B) = \sum_{t=1}^{s_2} P_2(Q_{j_t})$$

Провели n композиційних випробувань. Нехай $n_{B'}$ з них настала подія B' (в кожному випробуванні настала подія B), а в $n_{A'}$ з них настала подія A' , а в $n_{A'B'}$ з них настали події $A'B'$ одночасно.

$$\text{Тоді } W_n(A'/B') = \frac{n_{A'B'}}{n_{B'}}.$$

Кількісною мірою частоти є умовна ймовірність $P(A'/B')$.

Розглянемо композиційні випробування з множини $n_{B'}$. У кожному із них в другому випробуванні настала В. Так як перше та друге випробування не мають спільних випадкових факторів, то факт настання події В в другому випробуванні не накладає жодних умови на результат першого випробування.

Таким чином для першого випробування $n_{B'}$ випробувань є довільними. З них $n_{A'B'}$ випробувань настала подія А. Тоді:

$$\frac{n_{A'B'}}{n_{B'}} = W_{n_{B'}}(A).$$

Кількісна міра $W_{n_{B'}}(A) = P_1(A) = P(A')$.

Звідки $P(A'/B') = P(A') \Rightarrow A' \text{ і } B' - \text{незалежні.}$

ЛЕКЦІЯ 9

$$P(A' \cap B') = P(A') \cdot P(B') = P_1(A) \cdot P_2(B)$$

$$P(A \oplus B) = P_1(A) \cdot P_2(B)$$

Наслідки:

1) $A = E_i, \quad B = Q_j$

$$P(E_i Q_j) = P_1(E_i) \cdot P_2(Q_j)$$

2) Зрозуміла структура загального випадку двох незалежних подій A', B' в композиції двох незалежних випробувань, а саме:

- A' задається складною подією А в першому випробуванні, в другому випробуванні може статися що завгодно;
- B' задається складною подією В в другому випробуванні, в першому може статися будь-яка подія.

Примітка. Цей факт можна отримати формально, якщо має місце

$$P(E_i Q_j) = P_1(E_i) \cdot P_2(Q_j)$$

$$A' \cap B' = \bigcup_{l=1}^{S_1} \bigcup_{t=1}^{S_2} \{E_{i_l} Q_{j_t}\}$$

$$P(A' \cap B') = \sum_{l=1}^{S_1} \sum_{t=1}^{S_2} P_1(E_{i_l}) P_2(Q_{j_t}) = \sum_{l=1}^{S_1} P_1(E_{i_l}) \sum_{t=1}^{S_2} P_2(Q_{j_t})$$

3) Отриманий результат при розв'язку практичних задач формально пишеться невірно, а по своїй суті правильно.

Нехай ми маємо 1 і 2 випробування, які незалежні за 2 означенням.

1 випробування

2 випробування

$A \subset \Omega_1$

$B \subset \Omega_2$

Встановивши факт незалежності 1 і 2 випробувань за другим означенням (не мають спільних випадкових факторів) маємо рівність

$$P(A \cdot B) = P_1(A) \cdot P_2(B).$$

Правильне прочитання цієї формули:

Ймовірність того, що 1 і 2 випробування незалежні за 2 означенням, тоді ймовірність того, що в першому випробуванні настане подія А, а в другому – складна подія В дорівнює добутку ймовірностей $P_1(A) \cdot P_2(B)$.

Загальний випадок (композиція n випробувань)

Нехай здійснюється n випробувань із скінченими просторами елементарних подій

$$\Omega_i = \{ E_1^i, E_2^i, \dots, E_{m_i}^i \}, P_i(E_j^i) > 0, \sum_{j=1}^{m_i} P_i(E_j^i) = 1, i = \overline{1, n}.$$

Складними подіями у кожному i-му випробуванні є всілякі підмножини множини Ω_i , що визначає σ_i . Задані на кожному σ_i ймовірнісні міри $P_i(\cdot)$. І, таким чином, маємо n імовірнісних просторів $(\Omega_i, \sigma_i, P_i)$, $i = \overline{1, n}$.

Композиція n випробувань – випробування, що полягає в проведенні всіх n випробувань.

Це композиційне випробування задає наступний простір Ω елементарних подій

$$\Omega = \Omega_1 \oplus \Omega_2 \oplus \dots \oplus \Omega_n,$$

тобто у якості елементарних подій композиції беруться всілякі n-ки виду

$$(E_{j_1}^1, E_{j_2}^2, \dots, E_{j_n}^n); \quad j_1 = \overline{1, m_1}; \dots; j_n = \overline{1, m_n}.$$

Випробування називаються незалежними, якщо ймовірність кожної n-ки визначається співвідношенням

$$P(E_{j_1}^1 E_{j_2}^2 \dots E_{j_n}^n) = \prod_{i=1}^n P_i(E_{j_i}^i).$$

Випробування називаються незалежними, якщо будь-які два з них не мають спільних випадкових факторів.

Покажемо, що якщо виконується друге означення, то виконується і перше.

В побудованому просторі (Ω, σ, P) виділимо клас подій A , що називаються прямокутниками. Нехай $A_i \in \sigma_i, i = \overline{1, n}$. Прямокутник

$$A = A_1 \oplus A_2 \oplus \dots \oplus A_n$$

містить ті та й тільки ті n -ки $E = (E_{j_1}^1, E_{j_2}^2, \dots, E_{j_n}^n)$, у яких $E_{j_i}^i \in A_i, i = \overline{1, n}$.

З означення ймовірності для композиції незалежних випробувань (за першим означенням: $P(E_{j_1}^1 E_{j_2}^2 \dots E_{j_n}^n) = \prod_{i=1}^n P_i(E_{j_i}^i)$) виходить, що ймовірність прямокутника A дорівнює:

$$P(A) = \sum_{E \in A} P(E) = \sum_{E_{j_1}^1 \in A_1} \dots \sum_{E_{j_n}^n \in A_n} P_1(E_{j_1}^1) \dots P_n(E_{j_n}^n) = \sum_{E_{j_1}^1 \in A_1} P_1(E_{j_1}^1) \dots \sum_{E_{j_n}^n \in A_n} P_n(E_{j_n}^n) = P_1(A_1) P_2(A_2) \dots P_n(A_n). \quad (9.1)$$

Загальна структура незалежних подій в композиційному просторі задається прямокутниками виду

$$A'_i = \Omega_1 \oplus \Omega_2 \oplus \dots \oplus \Omega_{i-1} \oplus A_i \oplus \Omega_{i+1} \oplus \dots \oplus \Omega_n, \quad (9.2)$$

тобто подіями, які можна інтерпретувати наступним чином: в i -му випробуванні відбувається подія A_i , а в інших випробуваннях відбувається що завгодно.

Очевидно, що

$$P(A'_i) = P_i(A_i).$$

Оскільки $A = A_1 \oplus A_2 \oplus \dots \oplus A_n = \bigcap_{i=1}^n A'_i$, то

$$P\left(\bigcap_{i=1}^n A'_i\right) = P(A) = \prod_{i=1}^n P_i(A_i) = \prod_{i=1}^n P(A'_i) \quad (9.3)$$

тобто при різних i події A'_i незалежні.

Для доведення, що перше та друге означення незалежних випробувань є еквівалентними, використаємо метод математичної індукції.

Доведення загального випадку: нехай вірно для $n-1$, покажемо, що вірно для n (довести самостійно), проілюструємо на наступному частковому випадку.

Нехай вірно для 2, покажемо, що вірно для 3.

1	2	3
$A_1 \subset \Omega_1$	$A_2 \subset \Omega_2$	$A_3 \subset \Omega_3$

Розглянемо композицію $1 \oplus 2$ випробувань.

$$A'_1 = A_1 \oplus \Omega_2$$

$$A'_2 = \Omega_1 \oplus A_2$$

$$C = A'_1 \cap A'_2 = A_1 \oplus A_2$$

Будемо вважати перше випробування композицією 1 та 2 випробувань, а друге – 3 випробування.

$$\Omega'_1 = \Omega_1 \oplus \Omega_2$$

$$\Omega' = \Omega'_1 \oplus \Omega_3$$

$$C' = C \oplus \Omega_3$$

$$A'_3 = \Omega'_1 \oplus A_3 = \Omega_1 \oplus \Omega_2 \oplus A_3$$

Так як у 1, 2, 3 випробувань немає спільних випадкових факторів, то і у композиції $1'$ і 3 також немає спільних випадкових факторів. Отже, C' і A'_3 незалежні.

$$P(A_1 \oplus A_2 \oplus A_3) = P(C' \cap A'_3) = P(C') \cdot P(A'_3) = P(A'_1 \cap A'_2) P(A'_3) =$$

$$P(A'_1) P(A'_2) P(A'_3) = P_1(A_1) \cdot P_2(A_2) \cdot P_3(A_3).$$

Для $n=3$ отримали формулу (9.1) і як наслідок формули (9.2), (9.3).

$$\text{Звідки, якщо } A'_i = E_{j_i}^i \quad i = \overline{1, n}, \text{ то } P(E_{j_1}^1 E_{j_2}^2 \dots E_{j_n}^n) = \prod_{i=1}^n P_i(E_{j_i}^i).$$

Тобто перше та друге означення незалежних випробувань – еквівалентні.

Примітка. На практиці остання формула використовується наступним чином.

$$i = \overline{1, n}, \quad A_i \subset \Omega_i$$

$$P(A_1 \dots A_n) = \prod_{i=1}^n P(A_i)$$

Нехай n випробувань не мають спільних випадкових факторів, тоді ймовірність того, що в першому випробуванні настане подія A_1 , в другому - $A_2, \dots$, в n -му - A_n , дорівнює добутку ймовірностей настання цих подій.

Елементи теорії інформації

Розглянемо два ймовірнісних простори із скінченною кількістю елементарних подій (кожний задамо таблицею, у якої у верхньому рядку вказані елементарні події, у нижньому – ймовірності цих подій).

Розглянемо два випробування:

$$\begin{array}{cc} 1 & 2 \\ \left(\begin{array}{ccc} E_1 & E_2 & E_3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right) & \left(\begin{array}{cccc} Q_1 & Q_2 & Q_3 & Q_4 \\ 10^{-9} & 10^{-9} & 10^{-9} & 1-3 \cdot 10^{-9} \end{array} \right) \end{array}$$

У якого випробування, першого чи другого, більша неоднозначність результату випробування (до випробування)?

Очевидно, що ступінь невизначеності результату випробування (до випробування) у першому ймовірнісному просторі вище, ніж у другому. Дійсно, до випробування у першому просторі ні одній з подій не можна віддати перевагу, а у другому майже напевно відбудеться подія Q_4 .

Виникає питання: чи можна якимось чином міряти ступінь невизначеності, що властива ймовірнісному простору. Першим, хто запропонував вираз для міри невизначеності, був американський вчений К. Шеннон.

Нехай деякий ймовірнісний простір Ω задано таблицею

$$\left(\begin{array}{cccc} E_1 & E_2 & \dots & E_s \\ P_1 & P_2 & \dots & P_s \end{array} \right), \quad \sum_{i=1}^s P_i = 1.$$

К. Шеннон ввів наступну кількісну міру (ентропію) для вимірювання неоднозначності результату випробування (до випробування)

$$H(\Omega) = - \sum_{i=1}^s P_i \log_2 P_i,$$

при цьому приймається, що при $P_k = 0$ $P_k \log_2 P_k = 0$.

Ентропії притаманні такі простіші властивості.

- 1) Якщо ймовірнісний простір не має невизначеності, тобто якесь $P_i = 1$, а інші дорівнюють нулеві, то ентропія дорівнює нулеві (очевидно).
- 2) Коли елементарні події рівноймовірні, тобто $P_i = \frac{1}{s}$, $i = \overline{1, s}$, тоді ентропія приймає максимальне значення.

ЛЕКЦІЯ 10

Одиниця вимірювання ентропії – 1 біт, це ентропія випробування, яке полягає у підкидання монетки. Маємо дві рівноймовірні події.

$$\Omega = \begin{pmatrix} \Gamma & \text{p} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$H(\Omega) = -\log_2 \frac{1}{2} = -\log_2 2^{-1} = 1 \text{ (біт)}$$

Маємо 2 випробування :

1 $\Omega_1 = \begin{pmatrix} E_1 \dots E_{m_1} \\ p_1 \dots p_{m_1} \end{pmatrix}$ $p_i \geq 0 \quad \sum_{i=1}^{m_1} p_i = 1$	2 $\Omega_2 = \begin{pmatrix} Q_1 \dots Q_{m_2} \\ q_1 \dots q_{m_2} \end{pmatrix}$ $q_j \geq 0 \quad \sum_{j=1}^{m_2} q_j = 1$
---	---

Робимо композицію 1-го та 2-го випробувань.

$$\Omega = \Omega_1 \oplus \Omega_2 = \left\{ \begin{pmatrix} (E_i, Q_j) \\ P(E_i, Q_j) \end{pmatrix} \right\}, i = \overline{1, m_1}; j = \overline{1, m_2}$$

Якщо 1 і 2 випробування не мають спільних випадкових факторів, то

$$P(E_i Q_j) = p_i q_j.$$

Знайдемо ентропію:

$$H(\Omega) = - \sum_{i=1}^s P_i \log_2 P_i$$

$$H(\Omega) = - \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} P(E_i Q_j) \log_2 P(E_i Q_j)$$

Задача

Нехай p_i та q_j - фіксовані. Коли $H(\Omega)$ буде максимальною? Це буде відбуватися, коли випробування незалежні. А абсолютно максимальною, коли випробування незалежні і події рівноймовірні.

Нехай обидва випробування незалежні.

$$\begin{aligned} H(\Omega) &= - \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} P(E_i Q_j) \log_2 P(E_i Q_j) = \\ &= - \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} p_i q_j \log_2 (p_i q_j) \\ &= - \sum_{i=1}^{m_1} p_i \sum_{j=1}^{m_2} q_j \log_2 q_j - \sum_{j=1}^{m_2} q_j \sum_{i=1}^{m_1} p_i \log_2 p_i = H(\Omega_1) + H(\Omega_2). \end{aligned}$$

Примітки:

1. Логарифм добутку дорівнює сумі логарифмів.
2. З'являється дві суми, в яких змінюємо порядок індексування.
3. $\sum_{i=1}^{m_1} p_i = 1$, $\sum_{j=1}^{m_2} q_j = 1$.

Таким чином, для композиції двох випробувань маємо:

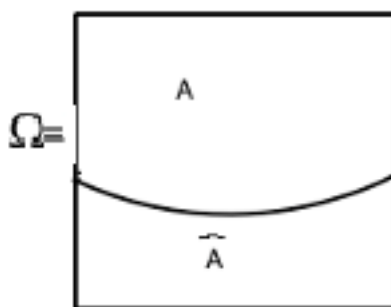
$$0 \leq H(\Omega_1 \oplus \Omega_2) \leq H(\Omega_1) + H(\Omega_2).$$

Біноміальний розподіл

Випробуваннями Бернуллі називаються n незалежних випробувань в кожному з яких може настати подія A чи подія $\bar{A}$ і в кожному з яких

$$P(A)=p, \quad P(\bar{A}) = 1 - p = q.$$

Зауваження. Події A і $\bar{A}$ не обов'язково є елементарними. Наприклад, всі точки прямокутника.



Задача:

Знайти ймовірність того, що внаслідок проведення n випробувань Бернуллі, подія A настане m разів ($m \leq n$).

Зробимо композицію n незалежних випробувань

Тоді $\Omega = \{A^{\alpha_1} A^{\alpha_2} \dots A^{\alpha_n}\}$. Кількість - 2^n .

$\alpha_i = \{0,1\}$, $A^0 = \bar{A}$, $A^1 = A$

Тоді, $P(A^{\alpha_1} \dots A^{\alpha_n}) = p^{\sum_{i=1}^n \alpha_i} q^{n - \sum_{i=1}^n \alpha_i}$.

Розглянемо в композиційному ймовірнісному просторі складну подію B : в n випробуваннях подія A відбулася m разів ($m \leq n$). Подія B складається з C_n^m елементарних подій, в кожному з яких символ A входить m разів, а символ $\bar{A}$ - $(n-m)$ разів.

Ймовірність кожної з цих елементарних подій однакова і дорівнює $p^m q^{n-m}$, незалежно від того, на яких позиціях стоїть подія A .

Тоді за означенням шукана ймовірність події B (тобто $P_n(m)$) дорівнює (за третьою аксіомою):

$$P_n(m) = C_n^m p^m q^{n-m},$$

$$\text{де } C_n^m = \frac{n!}{m!(n-m)!}.$$

Теорія ймовірності перестала бути наукою, яка формалізує азартні ігри, а стала математичним апаратом, що використовується фахівцем в будь-якій галузі з введенням поняття випадкової величини.

Випадкові величини

Випадкова величина (інженерна інтерпретація) задається випробуванням, простір елементарних подій яких є множиною чисел.

Примітка: результати випробувань вимірюються деякими приладами.

Випадкові величини бувають:

- Дискретні,
- Абсолютно неперервні,
- Сума дискретних і абсолютно неперервних (неперервні).

Майже у всіх випадках використовуються лише дискретні і абсолютно неперервні випадкові величини (перші два класи).

Примітка: Третій клас не розглядається, оскільки рідко зустрічається і для опису використовується апарат узагальнених функцій.

Випадкова величина X називається **дискретною**, якщо внаслідок випробування вона може набути значення з скінченної або зліченної множини можливих значень.

$\Omega = \{x_1 \dots x_s\}$ або $\Omega = \{x_1, \dots x_s, \dots\}$.

$$X \sim \begin{pmatrix} x_1 & \dots & x_s \\ p_1 & \dots & p_s \end{pmatrix}, \sum_{i=1}^s p(x_i) = 1$$

Неперервною випадковою величиною називається величина, простором елементарних подій якої може бути числова вісь, піввісь, відрізок числової осі, об'єднання відрізків числової осі, для якої існує функція розподілу та щільність, що належить класу неперервних функцій або кусково-неперервних функцій з обмеженою кількістю розривів чи стрибків 1-го роду.

Основною характеристикою, що описує всі випадкові величини є функція розподілу.

Функцією розподілу $F(x)$ називається числова скалярна функція дійсного аргументу, яка при довільному фіксованому значенні свого аргументу дорівнює ймовірності настання наступної події $P(X < x)$

Розглянемо $(X < x)$, x – довільне дійсне число.

$$X \sim \left(\begin{matrix} x_1 & \dots & x_s \\ p_1 & \dots & p_s \end{matrix} \right), \sum_{i=1}^s p(x_i) = 1.$$

В цьому випадку, ця подія. – об'єднання всіх елементарних подій, кожна з яких менша за x ($(X < x) = \bigcup_{x_i < x} \{x_i\}$).

Ймовірність настання цієї події означає, ймовірність того що внаслідок випробування настане будь-яка елементарна подія, менша, ніж фіксоване x .

У випадку неперервної випадкової величини $(X < x) = (-\infty, x) \cap \Omega$.

Примітка 1: В цьому випадку неможлива подія дорівнює $(-\infty, \infty)/\Omega$

Примітка 2: $\forall x P(X = x) \equiv 0$.

ЛЕКЦІЯ 11

Ймовірність настання в випробуванні будь-якого числа дорівнює нулю, а відмінна від нуля ймовірність настання – множина нескінченно-незлічених чисел (для складних подій).

$$F(x) = P(X < x)$$

$$\text{Дискретний випадок: } X < x \Rightarrow \bigcup_{\forall x_i} x_i < x \quad F(x) = P\left(\bigcup_{\forall x_i} x_i < x\right) = \sum_{\forall x_i < x} p_i$$

$$P\{a \leq X < b\} = P(\bigcup x_i \in [a, b) \cap \Omega).$$

Властивості функції розподілу:

$$1) F(x) = P(X < x)$$

$$F(-\infty) = 0$$

$$F(+\infty) = 1$$

$$\text{Примітка. } \exists \lim_{N \rightarrow \infty} F(-N)$$

$$\exists \lim_{N \rightarrow \infty} F(N)$$

$$a < b; (-\infty; b) = (-\infty, a) \cup [a, b)$$

$$(-\infty < X < b) = (-\infty < X < a) \cup [a \leq X < b)$$

$$(X < b) = (X < a) \cup [a \leq X < b)$$

$$P(X < b) = P(X < a) + P[a \leq X < b)$$

$$F(b) = F(a) + P([a \leq X < b))$$

З цього випливає дві відомі властивості функції розподілу:

- 2) Ймовірність того, що внаслідок випробування випадкова величина X прийме числове значення, що належить інтервалу $[a, b)$ дорівнює різниці значень функції розподілу на кінцях цього інтервалу

$$P([a \leq x < b)) = F(b) - F(a).$$

- 3) Для всіх $a < b \Rightarrow F(a) \leq F(b)$, тобто функція розподілу – неспадна.

- 4) Якщо випадкова величина X – неперервна, то функція розподілу – неперервна функція. Тому що ймовірність настання будь-якого числа тотожно дорівнює нулю, отже, функція розподілу не буде мати стрибків.

Функцією щільності випадкової величини X називають числову скалярну функцію аргументу x , що задовольняє інтегральній рівності :

$$F(x) = \int_{-\infty}^x f(u) du . \quad (11.1)$$

Випадкова величина зветься абсолютно неперервною, якщо функція щільності, що задовольняє цій рівності, належить класу неперервних функцій чи кусково-неперервних функцій, що мають обмежену кількість стрибків першого роду.

Властивості функції щільності :

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$
3. $P(a \leq X < b) = F(b) - F(a)$

$$\int_{-\infty}^b f(x) dx - \int_{-\infty}^a f(x) dx = \int_a^b f(x) dx$$

Якщо випадкова величина є абсолютно неперервною, то ймовірність попадання випадкової величини в півінтервал $[a; b)$ дорівнює

$$P(a \leq X < b) = \int_a^b f(x) dx.$$

Якщо випадкова величина X – неперервна, то

$$P(a \leq X < b) = P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b).$$

4. Нехай на $[a, b]$ функція щільності $f(x)$ неперервна. Беремо довільну внутрішню точку цього відрізка ($\forall x \in [a, b]$). Тоді має місце наступна рівність, що еквівалентна (11.1)

$$f(x) = \frac{dF(x)}{dx}.$$

***Строге означення випадкової величини

(Ω, σ, P) – ймовірнісний простір

Розглянемо числову скалярну функцію, аргументом якої є елементарна подія: $\varphi(w), w \in \Omega$.

З позиції інженерії функція $\varphi(w)$ описує прилад, що вимірює результат випробування. В цьому випадку треба вміти розв'язувати 3 задачі :

- I. $\forall x$ знайти $P(\forall w : \varphi(w) < x)$.
- II. $\forall [a, b)$ знайти $P(\forall w : a \leq \varphi(w) < b)$
- III. Знайти таку множину K , елементами якої є довільні підмножини чисел числової осі, що для кожної такої підмножини виконується:

$$B \in K, w \in \Omega \quad P(\forall w \quad \varphi(w) \in B) \text{ існує.}$$

Примітка. Необов'язково $\varphi(w)$ приймає всі можливі значення на $[a, b]$.

Приклад. X – дискретне, тобто $\varphi(w)$ приймає обмежену чи нескінченно-злічену множину значень. Всі w , при яких $x(w) \in [a, b]$ дорівнює всім w , при яких $x(w)$ приймає значення x_i , що належать відрізку $[a, b]$.

ЛЕКЦІЯ 12

Для того, щоб можна було розв'язати 3 задачі, треба, щоб $\varphi(w)$ мала певні властивості.

Числова скалярна функція $\varphi(w)$ називається **вимірною відносно σ -алгебри в (Ω, σ, P)** , якщо для будь-якого фіксованого x множина всіх w , для кожної з яких $\varphi(w) < x$, є елементом σ -алгебри, тобто має ймовірність наставання.

Випадковою величиною зветься $\varphi(w)$, яка є вимірною відносно σ -алгебри.

Якщо провести випробування, то настане якесь w , якому відповідає $\varphi(w) = x$. Це число зветься **реалізацією випадкової величини** в цьому випробуванні.

Надалі випадкові величини будемо позначати $\varphi(w)$, $X(w)$, $Y(w)$, $Z(w)$, або X , Y , Z (великі літери латинського алфавіту), коли по $\varphi(w)$ і по (Ω, σ, P) будемо новий ймовірнісний простір (R, B, P_φ) , де R – числова вісь, B – борелівська алгебра, P_φ – скінченна адитивна міра, а результат випробування є безпосереднє число (побудова (R, B, P_φ) по (Ω, σ, P) буде показано далі).

Інженерний зміст $\varphi(w)$ (властивості вимірності)

1. Якщо $\varphi(w)$ – вимірна, то яке б число x ми не зафіксували на числовій осі, гарантовано існує ймовірність (за Колмогоровим). того, що в наслідок випробування настане елементарна подія ω , така, що $\varphi(w) < x$
2. З означення випадкової величини випливає, що існує числова скалярна функція $F(x)$, яка для будь-якого фіксованого значення свого аргументу $F(x) = P(\forall \omega, \varphi(\omega) < x)$.

Ця функція і є **функцією розподілу випадкової величини**.

Властивості функції розподілу

1) $0 \leq F(x) \leq 1$

2) Розглянемо події:

$$A_1 = \{\forall \omega; \varphi(\omega) < b\}$$

$$A_2 = \{\forall \omega; \varphi(\omega) < a\}$$

$$A_3 = \{\forall \omega; a \leq \varphi(\omega) < b\}, a < b.$$

Події A_2 і A_3 – несумісні

$$A_1 = A_2 \cup A_3, \text{ оскільки } \varphi(\omega) - \text{скалярна.}$$

A_3 – має ймовірність наставання, тому, що належить σ – алгебрі, оскільки $A_3 = A_1 \setminus A_2$

Тоді

$$P(A_1)=P(A_2)+P(A_3)$$

$$F(b)=F(a)+P(\omega: a \leq \varphi(\omega) < b).$$

Таким чином, ймовірність потрапляння в напівінтервал дорівнює різниці значень функції розподілу на його граничних точках:

$$P(\omega: a \leq \varphi(\omega) < b)=F(b)-F(a)$$

3) Функція розподілу - неспадна функція

$$4) \text{ Розглянемо перетин } \bigcap_{n=1}^{\infty} [x - \frac{1}{n}; x) = \{\emptyset\}$$

$$A_n = \{\forall \omega \ x - \frac{1}{n} \leq \varphi(\omega) < x\}, \text{ тоді } P(\bigcap_{n=1}^{\infty} A_n)=0$$

В нашому випадку $A_{n+1} \subset A_n$.

Має місце наступна теорема.

Теорема: Якщо береться нескінченний перетин подій A_n і справедливе включення $A_{n+1} \subset A_n$, то $P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n) = 0$.

Цей факт буде доведено далі в теоремі про еквівалентність означення збіжності з ймовірністю одиниця.

$$\lim_{n \rightarrow \infty} (F(x) - F(x - \frac{1}{n})) = 0$$

$$F(x) - F(x-0) = 0$$

Тобто, отримали наступну властивість: Функція розподілу – неперервна зліва.

В деяких підручниках доводиться, що функція розподілу неперервна справа. В цьому випадку функція розподілу задається як

$$F(x) = P(\forall \omega, \varphi(\omega) \leq x)$$

5) Доведемо, що існують

$$F(-\infty) = \lim_{N \rightarrow \infty} F(-N) = 0, F(\infty) = \lim_{N \rightarrow \infty} F(N) = 1.$$

$$\text{Розглянемо, } A_n = \{\forall \omega \ n-1 \leq \varphi(\omega) < n\}.$$

$$\Omega = \bigcup_{n=-\infty}^{\infty} A_n, \text{ де } A_n - \text{попарно несумісні.}$$

$$\sum_{n=-\infty}^{\infty} P(A_n) = 1 \Rightarrow \exists \lim_{N \rightarrow \infty} \sum_{n=-N+1}^N P(A_n) =$$

$$= \lim_{N \rightarrow \infty} \sum_{n=-N+1}^N (F(n) - F(n-1)) = \lim_{N \rightarrow \infty} (F(N) - F(-N)) = 1.$$

Примітка: $\lim_{N \rightarrow \infty} \sum_{n=-N+1}^N P(A_n) = \lim_{N \rightarrow \infty} \sum_{n=-N+a}^N P(A_n)$, де a – довільне фіксоване число.

Оскільки, $0 \leq F(N) \leq 1$, $F(-N) \leq F(N)$, поєднати це можна тільки в одному випадку, коли існує $\lim_{N \rightarrow \infty} F(N) = 1$ і $\lim_{N \rightarrow \infty} F(-N) = 0$. Що й треба було довести.

ЛЕКЦІЯ 13

Довільна числова скалярна функція дійсного аргументу x має характеристичні властивості функції розподілу, якщо:

1. $F(-\infty) = 0$; $F(+\infty) = 1$
2. Неспадна
3. Неперервна зліва

Нагадаємо, для випадкових величин ми повинні розв'язати 3 задачі (перші дві – розв'язані):

1. $P\{\forall \omega: \varphi(\omega) < x\} = F(x)$, де $F(x)$ – функції розподілу;
2. $\forall a < b$
 $P(\forall \omega: a \leq \varphi(\omega) < b) = F(b) - F(a)$.

Розв'яжемо 3-ю задачу:

3. Знайти таку множину K , елементами якої є підмножини чисел числової осі, так що:

$$\{\forall \omega: \varphi(\omega) \in B\} \in \sigma.$$

Числова вісь $(-\infty; +\infty)$, інтервали $(-\infty; x) \forall x$ і будь-які півінтервали $[a, b)$, $a < b$ є елементами множини K

$$P\{\forall \omega, \varphi(\omega) < \infty\} = 1.$$

Покажемо, що множина K є σ -алгеброю, заданою на числовій осі.

Перевіряємо означення σ -алгебри.

1. $(-\infty; +\infty) \in K$
2. Нехай B належить K , тоді $\bar{B}$ теж повинно належати K , дійсно

$$\{\forall \omega: \varphi(\omega) \in \bar{B}\} = \overline{\{\forall \omega: \varphi(\omega) \in B\}} \in \sigma.$$

Оскільки $\{\forall \omega: \varphi(\omega) \in B\} \in \sigma$, то протилежна їй подія теж належить σ -алгебрі, звідси $\bar{B} \in K$.

3. Нехай $\forall B_i \in K, i \in I$. Розглянемо всі ω , для яких

$$\left\{ \forall \omega: \varphi(\omega) \in \bigcup_{i \in I} B_i \right\} = \bigcup_{i \in I} \{ \forall \omega, \varphi(\omega) \in B_i \}.$$

Так як для будь-якого B_i множина $w = \{ \forall \omega, \varphi(\omega) \in B_i \}$ є елементом σ -алгебри, то будь-яке об'єднання подій σ -алгебри є елементом σ -алгебри (за третьою властивістю σ -алгебри).

Довели, що K – σ -алгебра, що задана на числовій осі і містить всі півінтервали числової осі, а значить K містить мінімальну σ -алгебру, що містить всі півінтервали (борелівську σ -алгебру).

Отже, для довільної борелівської множини B означена ймовірність

$$P(\forall \omega: \varphi(\omega) \in B).$$

Ймовірність потрапляння випадкової величини φ в інтервал $[a, b)$ визначається через функцію розподілу за формулою

$$P(\forall \omega: a \leq \varphi(\omega) < b) = F(b) - F(a).$$

Довели, що існує функція розподілу $F(x)$ і побудували новий ймовірнісний простір $(R, \mathcal{B}, P_\varphi)$, причому, якщо елемент борелівської σ -алгебри – півінтервал, то його ймовірність відома і дорівнює різниці значень функції розподілу на кінцях цього півінтервалу. В цьому випадку $\Omega = R$.

Для нового побудованого ймовірнісного простору мають місце всі теореми теорії ймовірностей, але результатом випробування для нього є не w , а число на числовій осі (те, що і потрібно інженеру).

Наслідки.

Якщо випадкова величина дискретна, то $P(\forall \omega: \varphi(\omega) = x_i)$ для будь-якого x_i існує, тому що числа є елементами борелівської алгебри $\mathcal{B}$.

Далі буде показано, що $P(\forall \omega: \varphi(\omega) = x_i) = p_i = \lim_{n \rightarrow \infty} \frac{n_i}{n}$ (з ймовірністю 1).

Якщо випадкова величина неперервна, то Ω – нескінченна незлічена множина

$$P(\forall \omega, \varphi(\omega) = x) \equiv 0$$

Далі буде показано, що $P(\forall \omega: \varphi(\omega) < x) = F(x) = \lim_{n \rightarrow \infty} \frac{n_{x < x}}{n}$ (з ймовірністю 1)

$$P(\forall \omega, a \leq \varphi(\omega) < b) = F(b) - F(a)$$

– це ймовірність потрапляння у відрізок $[a, b)$.

І для дискретних, і для абсолютно неперервних випадкових величин функція розподілу дозволяє знайти імовірності всіх подій, які потрібні інженеру (тобто, не потрібно будувати $\varphi(\omega)$), але вона повинна бути вимірною відносно σ -алгебри.

Примітка. Знаючи, що $\varphi(\omega)$ породжує ймовірнісний простір $(R, \mathcal{B}, P_\varphi)$, то на ньому випадкові величини задаються великими літерами латинського алфавіту $X, Y, Z, \dots$, а для ймовірностей подій типу $P(\omega: X(\omega) \in B)$ часто будемо застосовувати скорочений запис $P(X \in B)$, а для функції розподілу випадкової величини X будемо використовувати запис $F(x) = P(X \leq x)$.

***Обернена теорема Колмогорова

(методологія доведення)

Розглянемо функцію $F(x)$ – довільну числову функцію, що задовольняє характеристичним властивостям функції розподілу: $F(-\infty) = 0$, $F(+\infty) = 1$, неспадна і неперервна зліва.

Розглянемо алгебру півінтервалів на числовій осі, на якій задамо скінченно-адитивну міру, а саме:

$$P([a, b)) = F(b) - F(a)$$

$$P\left(\bigcup_{i=1}^s [a_i, b_i)\right) = \sum_{i=1}^s (b_i - a_i), [a_i, b_i) \cap [a_j, b_j) = \emptyset, \forall i \neq j.$$

Ця міра є злічено-адитивною на алгебрі півінтервалів (як і в прикладі про голку). Тоді, за теоремою Каратеодорі, вона однозначно поширюється на мінімальну σ -алгебру, що містить півінтервали – борелівську алгебру.

Таким чином, на ймовірнісному просторі $((-\infty, \infty), \mathcal{B}, P_{F(x)})$ абстрактна функція розподілу $F(x)$ задала абстрактну випадкову величину X і її вигляд: $\varphi(x) = x$. Ми не знаємо, чи є таке випробування, що породжує цей простір, тому називаємо його абстрактним.

Розглянемо ймовірнісний простір (Ω, σ, P) , на якому задано випадкову величину $X(\omega)$, що є вимірною відносно σ .

Нехай $g(x)$ – числова скалярна функція дійсного аргументу x .

Виникає наступне питання. Чи є $g(X(\omega))$ випадковою величиною?

Функція $g(X(\omega))$ випадковою величиною лише тоді, коли вона буде вимірною відносно σ -алгебри.

Знайдемо умови, коли це виконується.

Функція $g(x)$ називається борелівською функцією, якщо для будь-якого дійсного y виконуються

$$(\forall x: g(x) < y) = B \in \mathcal{B}.$$

Теорема. Нехай $g(x)$ – борелівська функція, а $X(\omega)$ – випадкова величина, що задана на ймовірнісному просторі $(\Omega, \mathcal{B}, P)$. Тоді $Y(\omega) = g(X(\omega))$ – вимірна функція відносно σ -алгебри $\mathcal{B}$, тобто є випадковою величиною.

Доведення. Через те, що функція $g(x)$ борелівська, то для кожного y

$$B = \{x: g(x) < y\} \in \mathcal{B}.$$

Отже,

$\{\omega: Y(\omega) < y = g(X(\omega)) < y\} = \{\omega: X(\omega) \in B\} \in \sigma$ (оскільки борелівська алгебра належить $\mathcal{K}$). Таким чином, $Y(\omega)$ – випадкова величина.

Зазначимо, що клас борелівських функцій достатньо багатий: він містить, зокрема, всі аналітичні, неперервні та кусково-неперервні функції.

ЛЕКЦІЯ 14

Дискретні випадкові величини

Випадкова величини $X(\omega)$ на (Ω, σ, p) називається *дискретною*, якщо для будь-яких значень свого аргументу, вона може приймати одне значення з обмеженої кількості чисел чи нескінченно-зліченої кількості чисел.

$$(x_1 \dots x_s) \text{ або } (x_1 \dots x_s \dots).$$

Розглянемо всі $\forall \omega: X(\omega) = x_i; i = \overline{1, s \vee \infty}$.

Будь-яке число є елементом Борелівської алгебри, тож за попередньою теоремою ця множина всіх ω є елементом σ – алгебри та має ймовірність настання.

Таким чином дискретну випадкову величину можна задати таблицею:

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots & x_{s \vee \infty} \\ p_1 & p_2 & \dots & p_{s \vee \infty} \end{pmatrix}$$

Розглянемо функцію розподілу випадкової дискретної величини:

$$F(x) = P(X < x) = P(\cup \{x_i < x\}) = \sum_{\forall x_i < x} p_i.$$

Моменти(початкові та центральні) дискретних випадкових величин

Розглянемо дискретну випадкову величину задану таблицею:

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots & x_s \\ p_1 & p_2 & \dots & p_s \end{pmatrix}, \sum_{i=1}^s p_i = 1.$$

Проведемо n випробувань (незалежних):

$a_1, a_2, \dots, a_n$ де a_i – результат i -го випробування

Розглянемо середнє арифметичне: $\frac{1}{n} \sum_{i=1}^n a_n$

$$\frac{1}{n} \sum_{i=1}^n a_n$$

- це число відносно якого групуються результати проведення n випробувань над дискретною випадковою величиною X .

$$\frac{1}{n} \sum_{i=1}^n a_n = \sum_{i=1}^s x_i \frac{n_i}{n} = \sum_{i=1}^s x_i W_n(X = x_i),$$

де n_i – кількість випробувань, в кожному з яких настає x_i , W_n – частість його настання.

Якщо n достатньо велике, то для будь-якої довільної серії випробувань $W_n \approx P_i$, причому чим більше n , тим ближче W_n до p_i .

Математичним сподіванням або ймовірнісним середнім називається число на числовій осі $MX = \sum_{i=1}^s x_i p_i$, яке має наступний зміст: це число на числовій осі, відносно якого групуються результати конкретних випробувань над дискретною випадковою величиною в довільній достатньо великій серії повторів випробування.

Властивості математичних сподівань

1) Нехай кількість чисел в першому рядку таблиці – обмежена, тоді $MX = \sum_{i=1}^s x_i p_i$ – існує та обмежене по модулю.

Нехай кількість чисел в першому рядку таблиці – нескінченно-злічена, тоді $MX = \sum_{i=1}^{\infty} x_i p_i$ може бути:

- або числом обмеженим по модулю,
- або $\infty / -\infty$, коли існує границя.
- не існує (якщо суму можна представити у вигляді двох сум $+\infty$ і $-\infty$).

У подальшому детермінованою величиною (сталою) C будемо називати дискретну випадкову величину, що набуває єдине значення $x_1=C$ з імовірністю $p_1=1$.

$$C \sim \begin{pmatrix} x_1 = C \\ p_1 = 1 \end{pmatrix}.$$

2) Математичне сподівання сталої дорівнює цій сталій:

$$MC=C.$$

дійсно, за означенням

$$MC=x_1 p_1=C \cdot 1=C.$$

Розглянемо найпростіші функції від дискретної випадкової величини X за умови, що існує її математичне сподівання.

3) Математичне сподівання величини CX . Де C – стала, а X дискретна випадкова величина, дорівнює

$$MCX=CMX,$$

тобто стала виноситься з-під знаку математичного сподівання.

Таблиця для випадкової величини CX має вигляд

$$CX \sim \begin{pmatrix} Cx_1 & Cx_2 & \dots & Cx_s \\ p_1 & p_2 & \dots & p_s \end{pmatrix}.$$

За означенням математичного сподівання

$$MCX = \sum_{i=1}^s Cx_i p_i = C \sum_{i=1}^s x_i p_i = CMX.$$

4) Математичне сподівання суми $X+b$, де X дискретна випадкова величина, b – стала, дорівнює

$$M(X+b)=MX+b.$$

Будуємо таблицю

$$X+b \sim \begin{Bmatrix} x_1 + b & x_2 + b & \dots & x_s + b \\ p_1 & p_2 & \dots & p_s \end{Bmatrix}.$$

Маємо

$$M(X+b) = \sum_{i=1}^s (x_i + b) p_i = \sum_{i=1}^s x_i p_i + b \sum_{i=1}^s p_i = MX + b.$$

5) $M(aX+b)=aMX+b$, якщо a та b сталі, X – дискретна випадкова величина - довести по аналогії.

Нехай випадкова величина Y є функцією $f(X)$ від дискретної випадкової величини X .

Для випадкової величини $Y=f(X)$ будуємо таблицю

$$Y=f(X) \sim \begin{Bmatrix} f(x_1) & f(x_2) & \dots & f(x_s) \\ p_1 & p_2 & \dots & p_s \end{Bmatrix}.$$

Тут ми припускаємо, що усі $f(x_i)$ мають різні значення. Якщо це не так, об'єднуємо рівні значення функцій, сумуючи відповідні ймовірності. Математичне сподівання випадкової величини Y дорівнює:

$$MY = \sum_{i=1}^s f(x_i) p_i.$$

Приклад: Нехай $Y=X^2$, а випадкова величина X задана рядом розподілу.

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

Для того, щоб побудувати ряд розподілу для Y всі однакові числа потрібно об'єднати, а відповідні їм ймовірності – просумувати. Ряд розподілу матиме вигляд:

y_j	1	4	9	16
p_j	0,5	0,3	0,1	0,1

Тобто

$$MY = 1 \cdot 0,5 + 4 \cdot 0,3 + 9 \cdot 0,1 + 16 \cdot 0,1 =$$

$$= (-2)^2 \cdot 0,1 + (-1)^2 \cdot 0,2 + (1)^2 \cdot 0,3 + (2)^2 \cdot 0,2 + (3)^2 \cdot 0,1 + (4)^2 \cdot 0,1.$$

Примітка: Той самий результат отримаємо, коли $f(X)$ матиме стрибки першого роду. Але потрібно додатково визначити, чому дорівнює значення цієї функції в точці стрибка (лівій чи правій границі).

Початковим моментом k -го порядку випадкової величини X називається математичне сподівання випадкової величини $Y=X^k$:

$$v_k = MX^k = \sum_{i=1}^s x_i^k p_i.$$

Зрозуміло, що $v_1=MX$. Іноді позначають $MX=v$ (без індексу).

Центрована випадкова величина – це величина, що дорівнює $X'=X-MX$.

Математичне сподівання центрованої випадкової величини дорівнює нулю.

Дійсно, X – дискретна випадкова величина, і на підставі третьої властивості її математичного сподівання маємо

$$MX' = M(X-MX) = MX - MX = 0.$$

Центральним моментом k -го порядку випадкової величини X називається початковий момент k -го порядку випадкової величини X' :

$$\mu_k = M(X')^k = \sum_{i=1}^s (x_i - MX)^k p_i.$$

Дисперсією випадкової величини X називається її центральний момент другого порядку:

$$DX = \mu_2 = \sum_{i=1}^s (x_i - v)^2 p_i.$$

Додатне значення квадратного кореня з дисперсії має назву **середнього квадратичного відхилення** та позначається σ_X :

$$\sigma_X = +\sqrt{DX}.$$

ЛЕКЦІЯ 15

Якісний зміст дисперсії

Дисперсія – якісна міра ступеня групування результатів конкретних випробувань над дискретною випадковою величиною відносно її математичного сподівання в довільній серії випробувань. А саме, чим менше дисперсія – тим тісніше групуються конкретні результати випробувань відносно математичного сподівання.

Дійсно, нехай дисперсія мале число, це означає, що $\sum_{i=1}^s (x_i - v)^2 p_i$ – мала. А так як кожен доданок суми невід’ємний, то кожен доданок – дуже маленьке число.

Розглянемо довільний i -ий доданок, в якому $|x_i - v|$ – достатньо велике число, а, отже, квадрат цієї різниці ще більше число, а сам i -ий доданок дуже малий. Це можливо лише тоді, коли p_i – надзвичайно мале число, а, отже, x_i дуже рідко зустрічається в випробуванні. Таким чином, якщо дисперсія мала, то достатньо велику ймовірність наставання можуть мати лише ті x_i , в яких різниця $|x_i - v|$ – достатньо мале число. І чим менша дисперсія, тим менше ця різниця.

Математичні властивості дисперсії

1) Якщо $DX = 0$, то X – константа.

Доведення. Нехай усі x_i різні та

$$DX = \sum_{i=1}^s (x_i - v)^2 p_i = 0.$$

$$|x_i - v| > 0 \Rightarrow p_i = 0, \text{ але } \sum_{i=1}^s p_i = 1.$$

Тому що усі p_i не можуть дорівнювати нулеві (повинна виконуватись умова нормування), то ця рівність означає, що існує єдине $x_i = v$ і таке, що $p_i = 1$ (інших значень X може набувати лише з імовірністю нуль).

Примітка. Якщо X – константа, то $DX = 0$.

2) Стала виноситься з-під знаку дисперсії з квадратом:

$$D(CX) = C^2 D(X).$$

Примітка. Якщо X вимірюється в якихось одиницях, то дисперсія вимірюється в квадраті. Наприклад, X – м., DX – м².

Доведення. Дійсно,

$$Y = CX,$$

$$Y' = Y - MY = CX - M(CX) = CX - CMX = C(X - MX) = CX',$$

$$DY = M(Y')^2 = M(CX')^2 = MC^2(X')^2 = C^2M(X')^2 = C^2DX.$$

3) Дисперсія суми випадкової величини X та сталої C дорівнює дисперсії величини X :

$$D(X + C) = DX.$$

Доведення. Дійсно,

$$Y = X + C,$$

$$Y' = Y - MY = X + C - M(X + C) = X - MX = X',$$

$$DY = M(Y')^2 = M(X')^2 = DX.$$

4) Має місце наступне співвідношення між дисперсією і початковими моментами першого та другого порядку:

$$DX = v_2 - v_1^2.$$

Доведення. Дійсно,

$$DX = \sum_{i=1}^s (x_i - v_1)^2 p_i = \sum_{i=1}^s x_i^2 p_i - 2v_1 \sum_{i=1}^s x_i p_i + v_1^2 \sum_{i=1}^s p_i = v_2 - 2v_1^2 + v_1^2 = v_2 - v_1^2.$$

Твірна функція

Характеристичною функцією випадкової величини X називається функція

$$\boxed{\varphi_x(t) = Me^{itx}}$$

де i – уявна одиниця ($i = \sqrt{-1}$), t – дійсний аргумент.

$$\varphi_x(t) = Me^{itx} = \sum_{j=1}^{s \vee \infty} e^{itx_j} p_j$$

$$|\varphi_x(t)| \leq 1$$

Твірною функцією називається функція дійсної змінної t виду

$$m_x(t) = Me^{tx} = \sum_{j=1}^{s \vee \infty} e^{tx_i} p_i$$

Примітка. Твірна функція може не існувати, коли сума розглядається до ∞ . Будемо розглядати лише такі випадкові величини, для яких твірна функція є аналітичною функцією, тобто нескінченна сума для будь-якого t має границю. Якщо $\sum_{j=1}^s e^{tx_i} p_i$, то твірна функція є аналітичною функцією.

Властивості твірної функції

Нехай твірна функція є аналітичною функцією.

- 1) Похідна k -го порядку твірної функції у точці $t = 0$ дорівнює початковому моменту k -го порядку, тобто

$$\left. \frac{d^k m_x(t)}{dt^k} \right|_{t=0} = v_k.$$

Доведення. Дійсно, якщо твірна функція – аналітична, то k -ту похідну можна внести під знак суми

$$\left. \frac{d^k (\sum e^{x_i t} p_i)}{dt^k} \right|_{t=0} = \sum_{i=1}^s x_i^k e^{x_i t} p_i \Big|_{t=0} = \sum_{i=1}^s x_i^k p_i = v_k \quad (s \text{ може бути } +\infty).$$

- 2) Якщо твірна функція – аналітична, то її розвинення в ряд Маклорена має вигляд:

$$m_x(t) = 1 + \frac{v_1}{1!} t + \frac{v_2}{2!} t^2 + \dots + \frac{v_k}{k!} t^k + \dots$$

Примітка. Коефіцієнт при t^k дорівнює $\left. \frac{d^k (m_x(t))}{dt^k} \right|_{t=0}$.

ЛЕКЦІЯ 16

Приклад. Випадкова величина, що має біноміальний розподіл.

Маємо n випробувань Бернуллі, тоді ймовірність $P(A = m) = C_n^m p^m q^{n-m}$.

Дискретна випадкова величина X має біноміальний розподіл, якщо:

$$X \sim \left(\overline{m=0, n}, C_n^m p^m q^{n-m} \right), q = 1 - p.$$

$w = A^{\alpha_1} A^{\alpha_2} \dots A^{\alpha_n}$ (Ω складається з 2^n елементарних подій)

$\varphi(w) = m$, якщо в w подія A входить тільки m разів.

Знайдемо твірну функцію біноміального розподілу:

$$m_x(t) = \sum_{m=0}^n e^{tm} C_n^m p^m q^{n-m} = \sum_{m=0}^n C_n^m (e^t p)^m q^{n-m} = (e^t p + q)^n -$$

за формулою бінома Ньютона:

$$(a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}.$$

Неважко переконатися, що:

$$\left. \frac{dm_x(t)}{dt} \right|_{t=0} = np$$

$$\left. \frac{d^2 m_x(t)}{dt^2} \right|_{t=0} = n^2 p^2 + npq$$

За означенням початкового моменту 1-го порядку та 1-ю властивістю твірної функції:

$$MX = v_1 = np.$$

За 4-ю властивістю дисперсії, 1-ю властивістю твірної функції, а також попередньої рівності:

$$DX = v_2 - (v_1)^2 = n^2 p^2 + npq - n^2 p^2 = npq.$$

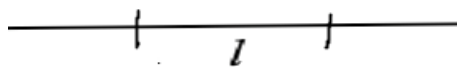
Розподіл Пуассона

Модель 1

Постановка задачі: Випробування – це проведення нескінченної кількості випробувань, результатом кожного з яких є поява числа на числової осі. В результаті випробування отримуємо нескінченно-злічену кількість чисел. Знайти ймовірність того, що на довільному відрізку довжини l числової осі з'явилося m чисел.

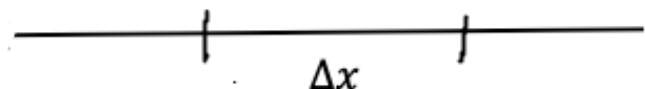
В такій постановці задачі вона не має розв'язку, тож на випадковий розподіл чисел числової осі накладаємо наступні умови:

1. **Стаціонарність** – ймовірність того, що на довільний відрізок фіксованої довжини попаде певна кількість чисел на числові осі, що залежить від довжини відрізка і не залежить від його розташування.



2. **Ординарність:**

1) ймовірність того, що на відрізок числової осі довжини Δx попаде число є нескінченною малою порядку Δx .



Тобто, $a\Delta x + o(\Delta x), 0 \leq a < \infty$,

($o(\Delta x)$ – нескінченно мала більш високого порядку, ніж Δx , якщо $\lim_{\Delta x \rightarrow 0} \frac{o(\Delta x)}{\Delta x} = 0$),

2) ймовірність появи двох чи більше чисел – це нескінченно мала більш високого порядку ніж Δx .

3. **Безпідсліддя:** ймовірність того, що на довільний відрізок попаде певна кількість чисел, не залежить від того, скільки чисел попаде на будь-які відрізки, що не перетинаються з ним (перетин пустий чи одне число)

Розв'язок задачі:

Беремо відрізок довжиною 1.



Зв'яжемо з ним випадкову величину X^1 - кількість чисел, що попали на відрізок довжиною 1.

$$X^1 \sim \left(\begin{matrix} i = \overline{0, \infty} \\ p_i^1 \end{matrix} \right)$$

$$MX^1 = \sum_{i=0}^{\infty} ip_i^1 = \lambda.$$

Нехай ми розглядаємо довільний відрізок довжини l . Зв'яжемо з ним X^l – кількість чисел, що попали на відрізок довжиною l , тоді

$$MX^l = \lambda \cdot l.$$

Для доведення цієї формули, використовуємо наступну формулу, що буде доведена далі:

$$M \sum_{i=1}^k X_i = \sum_{i=1}^k MX_i.$$

ЛЕКЦІЯ 17

1) Нехай l – натуральне число, тоді

$$X^l = \sum_{i=1}^l X_i^1 \Rightarrow MX^l = M \sum_{i=1}^l X_i^1 = \sum_{i=1}^l MX_i^1 = \lambda l.$$

2) Нехай l – раціональне число з обмеженою кількістю знаків після коми. Для спрощення розглянемо конкретний випадок. Наприклад, $l = k,137$ (k – ціле число)

$$\begin{aligned} X^l &= \sum_{i=1}^k X_i^1 + X^{0,137} \Rightarrow MX^l = k\lambda + MX^{0,137} \\ X^1 &= \sum_{i=1}^{1000} X_i^{0,001} \Rightarrow MX^1 = M \sum_{i=1}^{1000} X_i^{0,001} = \sum_{i=1}^{1000} MX_i^{0,001}; \\ MX_i^{0,001} &= \frac{\lambda}{1000} \\ MX^{0,137} &= \frac{\lambda \cdot 137}{1000} = \lambda \cdot 0,137 \end{aligned}$$

$$MX^l = k\lambda + \lambda \cdot 0,137 = \lambda \cdot l.$$

3) Нехай l – довільне число на числовій осі, задається десятковим дробом з нескінченною кількістю знаків після коми.

Так як формула вірна для чисел з довільним скінченним числом знаків після коми, то вірна і для нескінченності також.

Беремо довільний відрізок довжини l та розбиваємо на n відрізків однакової довжини.

$$\Delta x = \frac{l}{n} - \text{довжини відрізків.}$$

З кожним відрізком зв'язуємо випадкову величину $X^{\Delta x}$ – кількість точок, що потрапили в Δx .

В силу умови одинарності вважаємо, що

$$X^{\Delta x} \sim \begin{pmatrix} 0 & 1 \\ P_0^{\Delta x} & P_1^{\Delta x} \end{pmatrix}$$

де $P_0^{\Delta x}$ – ймовірність того, що на відрізок довжини Δx не попаде жодна точка;

$P_1^{\Delta x}$ – ймовірність того, що на відрізок довжини Δx попаде одна точка.

Примітка. Це припущення є нестрогим, тому і саме виведення є нестрогим (строге виведення формули Пуассона буде дано в спецкурсі).

Знайдемо математичне сподівання:

$$MX^{\Delta x} = \lambda \Delta x = P_1^{\Delta x}.$$

Таким чином, знайшли ймовірність попадання однієї точки на відрізок довжини Δx . Ймовірність того, що жодна точка не попаде на відрізок довжини Δx :

$$P_0^{\Delta x} = 1 - \lambda \Delta x.$$

За нашими припущеннями на відрізок довжини l попаде m точок тоді, коли до деяких m малих відрізків попаде по одній точці. За 3-ю властивістю та за біноміальним розподілом маємо

$$P_n(m) \approx C_n^m (\lambda \Delta x)^m (1 - \lambda \Delta x)^{n-m} = \frac{n!}{m!(n-m)!} \left(\frac{a}{n}\right)^m \frac{(1-a/n)^n}{(1-a/n)^m}$$

де $\Delta x = \frac{l}{n}$; $a = \lambda l$.

Точну ймовірність одержимо завдяки граничному переходу при $n \rightarrow \infty$ або $\Delta x \rightarrow 0$:

$$\pi_m = \lim_{n \rightarrow \infty} P_n(m) = \lim_{n \rightarrow \infty} \frac{n!}{m!(n-m)!} \left(\frac{a}{n}\right)^m \frac{\left(1 - \frac{a}{n}\right)^n}{\left(1 - \frac{a}{n}\right)^m} = \lim_{n \rightarrow \infty} \frac{(n-m+1) \cdots n}{n^m} \cdot \frac{a^m}{m!} \cdot \frac{\left(1 - \frac{a}{n}\right)^n}{\left(1 - \frac{a}{n}\right)^m} =$$

$$= \frac{a^m}{m!} e^{-a}, \quad \text{де } a = \lambda L.$$

Перевіримо, що $\sum_{m=0}^{\infty} \pi_m = 1$.

Дійсно, $e^{-a} \cdot \sum_{m=0}^{\infty} \frac{a^m}{m!} = e^{-a} \cdot e^a = 1$, оскільки $\sum_{m=0}^{\infty} \frac{a^m}{m!}$ – розвинення функції e^a в ряд

Маклорена.

Розподіл ймовірностей

$$\pi_m = \frac{a^m}{m!} \cdot e^{-a}, \quad m = \overline{0, \infty}, \quad \text{де } a > 0$$

називається розподілом Пуассона.

Відшукаємо твірну функцію для випадкової величини, що має розподіл Пуассона:

$$m_x(t) = \sum_{m=0}^{\infty} e^{mt} \frac{a^m}{m!} e^{-a} = e^{-a} \sum_{m=0}^{\infty} \frac{(ae^t)^m}{m!} = e^{-a} e^{ae^t} = e^{a(e^t - 1)}$$

для скінченних t .

Скориставшись властивостями твірної функції самостійно переконайтеся, що $v_1 = MX = a$, $v_2 = a^2 + a$, $DX = a$.

Розподіл Пуассона

Модель 2А

Постановка задачі. Маємо n випробувань Бернуллі, тоді ймовірність того, що в n випробуваннях подія A настане m разів відома і дорівнює

$$C_n^m p^m (1 - p)^{n-m}.$$

Нехай n – велике число, а p – достатньо мале. Така модель має назву закон рідких явищ.

Вводимо $a = np$, тоді точна формула

$$C_n^m \left(\frac{a}{n}\right)^m \left(1 - \frac{a}{n}\right)^{n-m};$$

При фіксованому a переходячи до границі при $n \rightarrow \infty$, отримаємо

$$\frac{a^m}{m!} e^{-a}.$$

Якщо n – достатньо велике число, а $\frac{a}{n}$ – достатньо мале, то точна ймовірність знаходиться в малому околі границі $\frac{a^m}{m!} e^{-a}$, яку і можна вважати наближеним значенням точної ймовірності.

Розподіл Пуассона

Модель 2В

Постановка задачі та ж сама, що і для моделі 2А, але ймовірність p настання події A залежить від кількості випробувань ($p = \frac{a}{n}$).

Кількість випробувань $n \rightarrow \infty$, тоді ймовірність настання події A m разів в нескінченній кількості випробувань дорівнює

$$\frac{a^m}{m!} e^{-a} \text{ – формула Пуассона.}$$

Неперервні випадкові величини

Якщо випадкові величини – неперервні (Ω – нескінченно-незлічена множина чисел), то ймовірність настання будь-якого числа дорівнює нулю. Тоді, можуть бути події, які теоретично можуть настати і мають ймовірність нуль, можуть бути події, які можуть теоретично не настати і мають ймовірність одиниця.

Покажемо на прикладах, якщо подія має ймовірність, що дорівнює нулю, вона ніколи не настає в обмеженій кількості випробувань і навпаки. Тобто для інженера події з ймовірністю одиниця є достовірними, а з ймовірністю нуль – неможливими.

ЛЕКЦІЯ 18

Приклад1: Випробування – композиція нескінченної кількості незалежних підкидань монети.

Елементарна подія – нескінченна послідовність О (орел) і Р (решка).

Ймовірність настання будь-якої події: $\lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n = 0$

З теорії множин відомо, що множина всіх різних нескінченних послідовностей Ω та P є нескінченно-незліченою.

Таким чином, Ω – нескінченно незлічена множина, а $P(\forall w)=0$.

Зафіксуємо будь-яку елементарну подію w – нескінченну послідовність O і P і починаємо випробування. Гарантовано, що на якомусь підкиданні літери не співпадуть. Ще раз повторимо випробування – знову отримаємо іншу w . Таким чином в обмеженій кількості композиційних випробувань зафіксована w не буде отримана.

Приклад 2: Внаслідок випробування з'являється число на числовій осі (десятковий дріб з нескінченною кількістю знаків після коми). Фіксуємо велику, але обмежену кількість знаків після коми. Ціла частина та фіксована після коми – це числа достатньо малого околу якогось числа. Якщо окіл достатньо малий, то ймовірність появи будь-якого наступного числа в будь-якому наступному розряді – рівноймовірна подія, з ймовірністю $\frac{1}{10}$.

Маємо ситуацію з попереднього прикладу, тільки замість монетки підкидаємо 10 граней кубик.

Неперервні випадкові величини

Простір елементарної події Ω – вся числова вісь, або відрізки числової осі (нескінченно-незлічена множина чисел).

ТЕОРЕМА

Нехай ймовірність появи будь-якої події (числа) числової осі $P(X=x) \equiv 0$, тоді її функція розподілу неперервна і навпаки, якщо функція розподілу неперервна, то $P(X=x) \equiv 0$.

ДОВЕДЕННЯ

$x = \bigcap_{n=1}^{\infty} [x, x + \frac{1}{n})$, $P(\bigcap_{n=1}^{\infty} A_n)$, де $A_{n+1} \subset A_n$, $A_n = (\forall w: \varphi(w) \in [x, x + \frac{1}{n})$.

Маємо:

$P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n) = P(\varphi(w)=x) = P(X=x)$.

$$P(X=x) = \lim_{n \rightarrow \infty} (F(x+\frac{1}{n}) - F(x)) = F(x+0) - F(x).$$

Якщо $P(X=x)=0$, то $F(x+0)=F(x)$,

то функція розподілу неперервна, оскільки ми довели її неперервність справа, а раніше була доведена її неперервність зліва.

Нехай функція розподілу неперервна. Тоді вона неперервна справа: $F(x+0)=F(x)$, тоді $P(X=x)=0$. Теорему доведено.

Випадкова величина **неперервна**, якщо її функція розподілу – неперервна.

Випадкова величина X зветься **абсолютно неперервною**, якщо для неї існує числова скалярна $f(x)$ з класу неперервних функцій чи кусково-неперервних функцій з обмеженою кількістю стрибків 1-го роду, що задовольняє наступній інтегральній рівності.

$$F(x) = \int_{-\infty}^x f(u) du \quad (18.1)$$

Функція $f(x)$, для якої справедлива рівність (18.1), називатиметься **функцією щільності**.

Примітка: Надалі розглядатимемо дискретні та абсолютно неперервні випадкові величини, які називатимемо просто неперервними.

Властивості:

- 1) $f(x) \geq 0$ (функція щільності невід'ємна), тому що $F(x)$ – неспадна
- 2) $P(X \in [a, b]) = \int_a^b f(x) dx$.

Доведення:

$$F(b) - F(a) = \int_{-\infty}^b f(x) dx - \int_{-\infty}^a f(x) dx = \int_a^b f(x) dx$$

$$\text{Наслідок: } \int_{-\infty}^{\infty} f(x) dx = 1$$

Будь-яка числова скалярна функція $f(x)$ дійсного аргументу має характерні властивості функції щільності, якщо:

- Вона неперервна чи кусково-неперервна з обмеженою кількістю стрибків 1-го роду

- Вона невід’ємна
- $\int_{-\infty}^{\infty} f(x)dx = 1$

Наслідок: Така функція задає абстрактну випадкову величину. Тобто ми не знаємо випробування, яке її породжує. Дійсно, абстрактна функція щільності задає абстрактну функцію розподілу ($F(x) = \int_{-\infty}^x f(u)du$).

3) Нехай на відрізку $(a; b)$ функція щільності $f(x)$ неперервна, то

$$\forall x \in (a; b) \Rightarrow \frac{dF(x)}{dx} = f(x). \quad (18.2)$$

Наслідок:

$$f(x) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x \geq 0}} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x \geq 0}} \frac{P(x \leq X \leq x + \Delta x)}{\Delta x}$$

$$P(x \leq X \leq x + \Delta x) = f(x)\Delta x + o(\Delta x).$$

Примітка: З математичного аналізу випливає наступне:

$$P(x \leq X \leq x + \Delta x) = f(x + \theta \Delta x)\Delta x + o(\Delta x); \forall \theta \in (0; 1).$$

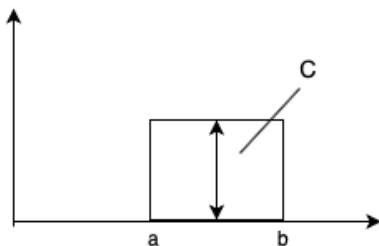
ЛЕКЦІЯ 19

Приклади абстрактних неперервних випадкових величин

1. РІВНОМІРНИЙ РОЗПОДІЛ

Неперервна випадкова величина X має рівномірний розподіл, якщо її абстрактна функція щільності дорівнює

$$f(x) = \begin{cases} 0, & x < a \\ c = \text{const}, & x \in [a, b] \\ 0, & x > b \end{cases}$$

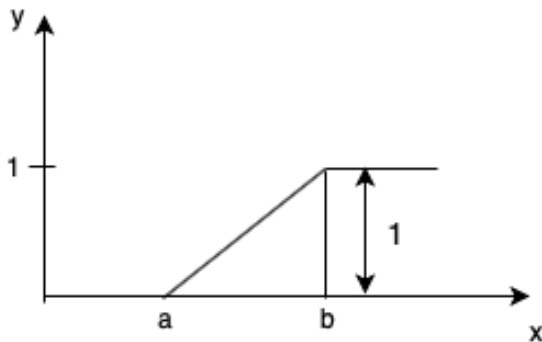


Функція $f(x)$ має характеристичні властивості функції щільності, отже

$$\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_a^b c dx = 1 \Rightarrow c = \frac{1}{b-a}.$$

Відповідно, абстрактна функція розподілу випадкової величини має вигляд

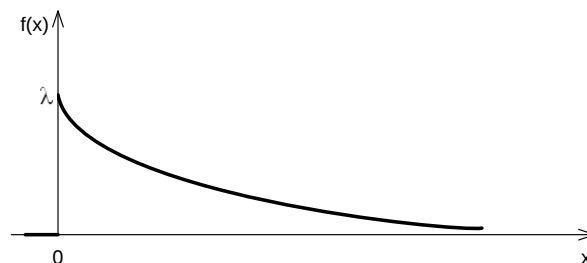
$$F(x) = \int_{-\infty}^x f(x)dx = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & x \in [a, b] \\ 1, & x > b \end{cases}$$



2. ЕКСПОНЕНЦІЙНИЙ РОЗПОДІЛ

Абстрактна випадкова величина X має експоненційний розподіл, якщо її функція щільності має вигляд

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}, \text{ де } \lambda > 0, \lambda - \text{константа.}$$



Очевидно, має місце

$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^{+\infty} \lambda e^{-\lambda x} dx = \int_0^{+\infty} d(-e^{-\lambda x}) = -e^{-\lambda x} \Big|_0^{\infty} = 1;$$

$$F(x) = \int_{-\infty}^x f(u)du = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Неперервна випадкова величина є математичною абстракцією і у чистому вигляді на практиці не зустрічається хоча б тому, що теоретично не може існувати вимірювальний пристрій, що обчислює реалізацію цієї величини. Тому дослідник

завжди має справу з дискретними випадковими величинами (результатами вимірювання). На практиці дискретну випадкову величину апроксимують неперервною, коли множина значень величини рівномірно і „щільно” заповнює деякий відрізок числової осі. Це зручно, бо ймовірнісні характеристики задаються не за допомогою сукупності чисел – ймовірностей елементарних подій, а за допомогою функції $F(x)$ або $f(x)$.

У теоретичному плані багато конструктивних результатів одержано за допомогою поняття неперервної випадкової величини, зокрема, граничні теореми.

Ймовірнісні характеристики неперервних випадкових величин

Нехай X – неперервна випадкова величина, задана функцією щільності $f(x)$.

Розглянемо довільну скалярну числову функцію $\varphi(x)$ дійсного аргументу з класу неперервних або кусково-неперервних функцій з обмеженою кількістю стрибків 1-го роду.

Розглянемо випадкову величину $\varphi(X)$. Тоді математичне сподівання нової випадкової величини $\varphi(X)$ має вигляд

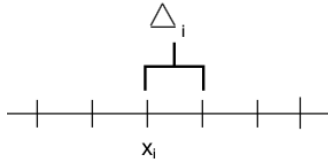
$$M\varphi(X) = \int_{-\infty}^{\infty} \varphi(x)f(x)dx. \quad (19.1)$$

Приклад. $\varphi(x) = x^2 \Rightarrow MX^2 = \int_{-\infty}^{\infty} x^2 f(x)dx.$

Обґрунтування формули (19.1)

Примітка 1. Для спрощення викладу будемо вважати, що $\varphi(x)$ і $f(x)$ – неперервні функції. В загальному випадку треба числову вісь розбити на відрізки, на кожному з яких ці дві функції $\varphi(x), f(x)$ є неперервними та для кожного відрізка повторити обґрунтування, які будуть розглянуті нижче.

Примітка 2. Треба домовитись, чому дорівнюють функції $\varphi(x)$ і $f(x)$ в точках стрибка (границі зліва та справа).



Числову вісь розбиваємо на нескінченно-злічену кількість відрізків однакової малої довжини Δx . Позначимо через x_i лівий кінець відрізка Δi -го відрізка, $i = \overline{-\infty; \infty}$.

Δx – мале, щоб можна було застосувати третю властивість функції щільності.

Неперервну випадкову величину X *апроксимуємо* дискретною випадковою величиною X^* наступним чином:

Якщо X внаслідок випробування попало в Δi -ий відрізок, то в цьому випробуванні дискретна випадкова величина X^* приймає значення x_i (лівий кінець Δi -го відрізка).

Чим менше Δx , то тим точніше X^* апроксимує X .

Побудуємо таблицю для випадкової величини X^*

$$X^* \sim \left(\begin{array}{c} x_i, \quad i = \overline{-\infty, \infty} \\ f(x_i)\Delta x + o(\Delta x) \end{array} \right)$$

Неперервну випадкову величину $Y = \varphi(X)$ апроксимуємо дискретною випадковою величиною $Y^* = \varphi(X^*)$. Оскільки функція $\varphi(x)$ – неперервна, то чим менше Δx , тим точніше Y^* апроксимує Y .

Використаємо загальну формулу

$$M\varphi(X) = \sum_{i=1}^{SV\infty} \varphi(x_i)p_i,$$

де X – дискретна випадкова величина, тоді

$$M\varphi(X^*) = \sum_{i=-\infty}^{\infty} \varphi(x_i)[f(x_i)\Delta x + o(\Delta x)].$$

Щоб знайти $M\varphi(X)$, треба $\Delta x \rightarrow 0$.

Якщо $|\int_{-\infty}^{\infty} \varphi(x)f(x)dx|$ існує і обмежений, то

$$\lim_{\Delta x \rightarrow 0} M\varphi(X^*) = \lim_{\Delta x \rightarrow 0} \sum_{i=-\infty}^{\infty} \varphi(x_i)(f(x_i)\Delta x + o(\Delta x)) = \int_{-\infty}^{\infty} \varphi(x)f(x)dx = M\varphi(X).$$

Звідси випливає, що

$$MX = \int_{-\infty}^{\infty} xf(x)dx \quad (19.2)$$

$$\varphi(x) = x$$

Так як формула (19.2) була обґрунтована як граничний перехід від математичного сподівання дискретної випадкової величини, то математичне сподівання неперервної випадкової величини має той самий фізичний зміст як і дискретна випадкова величина – це точка на числовій осі, відносно якої групуються результати випробування над неперервною випадковою величиною в достатньо великій серії випробувань.

Довести самостійно, що математичне сподівання неперервної випадкової величини має ті самі властивості, що і для дискретної випадкової величини

$$M(aX+b)=aMX+b.$$

Для неперервної випадкової величини початкові моменти, центральні моменти, дисперсія та середнє квадратичне відхилення означаються співвідношеннями:

$$\nu_k = MX^k = \int_{-\infty}^{+\infty} x^k f(x) dx,$$

$$\mu_k = \int_{-\infty}^{+\infty} (x - \nu_1)^k f(x) dx,$$

$$DX = \mu_2 = \int_{-\infty}^{+\infty} (x - \nu_1)^2 f(x) dx,$$

$$\sigma_X = +\sqrt{DX}.$$

Фізичний зміст дисперсії неперервної випадкової величини такий же самий, як і у дискретної випадкової величини – якісна міра ступеня групування конкретних результатів випробування над неперервною випадковою величиною, а саме, чим менше дисперсія – тим тісніше результати групуються відносно математичного сподівання.

Самостійно переконатися, що

$$D(CX) = C^2 DX,$$

$$D(C + X) = DX,$$

$$DX = \nu_2 - \nu_1^2.$$

Аналогічно дискретному випадку твірна функція неперервної випадкової величини означається як

$$m_X(t) = Me^{Xt} = \int_{-\infty}^{+\infty} e^{xt} f(x) dx.$$

У припущенні аналітичності функції $m_X(t)$ довести самостійно, що

$$\left. \frac{d^k m_X(t)}{dt^k} \right|_{t=0} = v_k,$$

$$m_X(t) = 1 + \frac{v_1}{1!} t + \frac{v_2}{2!} t^2 + \dots + \frac{v_k}{k!} t^k + \dots$$

Абстрактна неперервна випадкова величина, що має нормальний розподіл

Абстрактна неперервна випадкова величина має нормальний розподіл, якщо її абстрактна функція щільності дорівнює

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad \sigma > 0, \quad x \in (-\infty, +\infty).$$

Перевіримо, чи виконуються характеристичні властивості для функції щільності:

- 1) $f(x)$ – неперервна.
- 2) $f(x)$ – невід’ємна.
- 3) Перевіримо умову нормування:

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left| \begin{array}{l} z = \frac{x-a}{\sigma} \\ dx = \sigma dz \end{array} \right| = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz.$$

Перевірити самостійно, що $\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}$ (інтеграл Пуассона). Звідси буде

виходити, що $\int_{-\infty}^{+\infty} f(x) dx = 1$.

ЛЕКЦІЯ 20

Знайдемо твірну функцію нормального розподілу:

$$m_x(t) = Me^{tX} = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-a)^2}{2\sigma^2}} dx =$$

(зробимо заміну змінних:

$$\begin{aligned} \frac{x-a}{\sigma} &= z, \\ x &= \sigma z + a, \\ dx &= \sigma dz. \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma z + a) - \frac{z^2}{2}} dz =$$

$$e^{at} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2t\sigma z + t^2\sigma^2 - t^2\sigma^2)} dz = e^{(at + \frac{t^2\sigma^2}{2})} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t\sigma)^2}{2}} dz.$$

Примітка: Показати самим, що

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t\sigma)^2}{2}} dz = 1.$$

$$m_x(t) = e^{at + \frac{t^2\sigma^2}{2}} - \text{твірна функція нормального розподілу}$$

Так як для нормального розподілу твірна функція існує і є аналітичною, тоді вона задовольняє двом властивостям, що були виведені раніше:

$$\left. \frac{dm_x(t)}{dt} \right|_{t=0} = a = v_1 = v,$$

$$\left. \frac{d^2 m_x(t)}{dt^2} \right|_{t=0} = v_2 = a^2 + \sigma^2 \Rightarrow DX' = v_2 - v_1^2 = \sigma^2.$$

Функція щільності нормального розподілу:

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-v)^2}{2\sigma^2}} = n(x, v, \sigma).$$

Функція розподілу:

$$F(x) = N(x, v, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-v)^2}{2\sigma^2}} du.$$

$$m_x(t) = e^{vt + \frac{t^2\sigma^2}{2}} - \text{твірна функція нормального розподілу.}$$

$$n(x, 0, \sigma) \Rightarrow MX = 0.$$

$e^a = \sum_{m=0}^{\infty} \frac{a^m}{m!}$ – ряд Маклорена.

$m_x(t) = 1 + \frac{v_1}{1!}t + \frac{v_2}{2!}t^2 + \frac{v_3}{3!}t^3 + \dots$ - представлення твірної функції у вигляді ряду Маклорена.

Теорема.

Якщо $MX = 0 \Rightarrow$ всі непарні початкові моменти дорівнюють нулю, довести самим.

Функція Лапласа

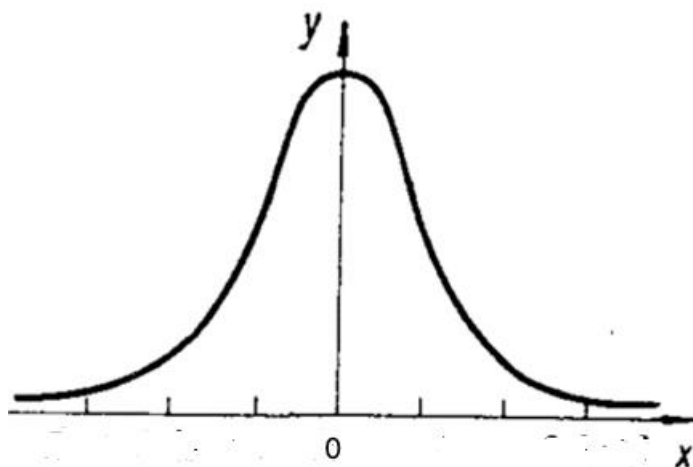
Обчислимо ймовірність попадання випадкової величини X в інтервал $[a;b]$:

$$P(a \leq X \leq b) = F(b) - F(a) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Значення цього інтегралу можна знайти лише чисельними методами.

Функцією Лапласа зветься числова скалярна функція:

$\Phi_0(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{u^2}{2}} du$ – інтеграл від функції щільності нормованого нормального розподілу.



Властивості:

1. $\Phi_0(-\infty) = -\frac{1}{2}, \Phi_0(+\infty) = \frac{1}{2}$.
2. $\Phi_0(-z) = -\Phi_0(z)$ – непарна.
3. $\Phi_0(z)$ – табульована.

Тобто для невід’ємних z із заданим кроком, із заданою похибкою функцію Лапласа обчислюють чисельними методами. На практиці, для конкретного значення аргументу шукаємо відповідне значення функції в таблиці.

z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
0,00	0,0000	0,50	0,1915	1,00	0,3413	1,50	0,4332	2,00	0,4772	2,50	0,49865
0,01	0,0040	0,51	0,1950	1,01	0,3438	1,51	0,4345	2,01	0,4783	2,51	0,49931
0,02	0,0080	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,02	0,4793	2,52	0,49996
0,03	0,0120	0,53	0,2019	1,03	0,3485	1,53	0,4370	2,03	0,4803	2,53	0,499941
0,04	0,0160	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,04	0,4812	2,54	0,499928
0,05	0,0199	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,05	0,4821	2,55	0,499968
0,06	0,0239	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,06	0,4830	2,56	0,499997
0,07	0,0279	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,07	0,4838	2,57	0,499997
0,08	0,0319	0,58	0,2190	1,08	0,3599	1,58	0,4429	2,08	0,4846		
0,09	0,0359	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,09	0,4854		
0,10	0,0399	0,60	0,2257	1,10	0,3643	1,60	0,4452	2,10	0,4861		
0,11	0,0438	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,11	0,4868		
0,12	0,0478	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,12	0,4875		
0,13	0,0517	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,13	0,4881		
0,14	0,0557	0,64	0,2389	1,14	0,3728	1,64	0,4495	2,14	0,4887		
0,15	0,0596	0,65	0,2422	1,15	0,3749	1,65	0,4506	2,15	0,4893		
0,16	0,0636	0,66	0,2454	1,16	0,3770	1,66	0,4516	2,16	0,4898		
0,17	0,0675	0,67	0,2486	1,17	0,3790	1,67	0,4525	2,17	0,4904		
0,18	0,0714	0,68	0,2517	1,18	0,3810	1,68	0,4535	2,18	0,4909		
0,19	0,0753	0,69	0,2549	1,19	0,3830	1,69	0,4545	2,19	0,4913		
0,20	0,0793	0,70	0,2580	1,20	0,3849	1,70	0,4554	2,20	0,4918		
0,21	0,0832	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,21	0,4922		
0,22	0,0871	0,72	0,2642	1,22	0,3889	1,72	0,4573	2,22	0,4927		
0,23	0,0910	0,73	0,2673	1,23	0,3907	1,73	0,4582	2,23	0,4931		
0,24	0,0948	0,74	0,2703	1,24	0,3925	1,74	0,4591	2,24	0,4934		
0,25	0,0987	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,25	0,4938		
0,26	0,1026	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,26	0,4941		
0,27	0,1064	0,77	0,2794	1,27	0,3980	1,77	0,4616	2,27	0,4945		
0,28	0,1103	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,28	0,4948		
0,29	0,1141	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,29	0,4951		
0,30	0,1179	0,80	0,2881	1,30	0,4032	1,80	0,4641	2,30	0,4953		
0,31	0,1217	0,81	0,2910	1,31	0,4049	1,81	0,4649	2,31	0,4956		
0,32	0,1255	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,32	0,4959		
0,33	0,1293	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,33	0,4961		
0,34	0,1331	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,34	0,4963		
0,35	0,1368	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,35	0,4965		
0,36	0,1406	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,36	0,4967		
0,37	0,1443	0,87	0,3078	1,37	0,4147	1,87	0,4693	2,37	0,4969		
0,38	0,1480	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,38	0,4971		
0,39	0,1517	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,39	0,4973		
0,40	0,1554	0,90	0,3159	1,40	0,4192	1,90	0,4713	2,40	0,4974		
0,41	0,1591	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,41	0,4976		
0,42	0,1628	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,42	0,4977		
0,43	0,1664	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,43	0,4979		
0,44	0,1700	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,44	0,4980		
0,45	0,1736	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,45	0,4981		
0,46	0,1772	0,96	0,3315	1,46	0,4279	1,96	0,4750	2,46	0,4982		
0,47	0,1808	0,97	0,3340	1,47	0,4292	1,97	0,4756	2,47	0,4984		
0,48	0,1844	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,48	0,4985		
0,49	0,1879	0,99	0,3389	1,49	0,4319	1,99	0,4767	2,49	0,4986		

Повертаємось до нашої задачі.

$$P(a \leq X \leq b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\nu)^2}{2\sigma^2}} dx =$$

Зробимо заміну:

$$\begin{aligned} \frac{x - \nu}{\sigma} &= z, \\ x &= \sigma z + \nu, \\ dx &= \sigma dz. \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\nu}{\sigma}}^{\frac{b-\nu}{\sigma}} e^{-\frac{z^2}{2}} dz = \Phi_0\left(\frac{b-\nu}{\sigma}\right) - \Phi_0\left(\frac{a-\nu}{\sigma}\right).$$

Задача: Нехай X – неперервна випадкова величина, $f(x)$ – функція щільності для X , $h(x)$ – довільна неперервна числова скалярна функція для дійсного аргументу x . Знайти функцію щільності випадкової величини $H=h(X)$ ($g_H(h)$).

Приклад:

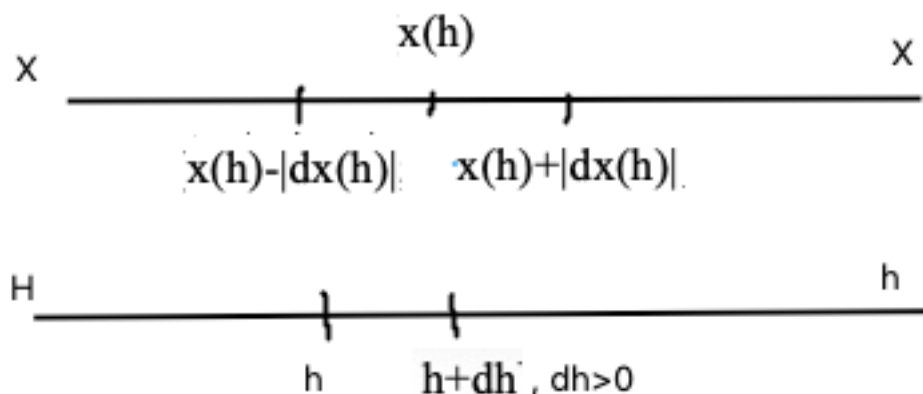
Дано: $h(x)=x^2$, знайти функцію щільності випадкової величини X^2 .

Частковий випадок

$h(x) \rightarrow x(h)$ — функція $h(x)$ має єдину обернену функцію $x(h)$.

$$h(x) \rightarrow \begin{cases} x_1(h) \\ \dots \\ x_k(h) \end{cases} \text{ — загальний випадок - } k \text{ обернених функцій.}$$

Розглянемо задачу для однієї оберненої функції:



$$P(h \leq H \leq h + dh) = g_H(h)dh + o(dh),$$

де $g_H(h)$ — невідома функція щільності випадкової величини H .

Якщо H попадає у відрізок $[h, h+dh]$, то X попадає у:

відрізок $[x(h); x(h)+|dx(h)|]$, якщо $dx(h)$ додатній.

відрізок $[x(h)-|dx(h)|; x(h)]$, якщо $dx(h)$ від'ємний.

Незалежно від того, в який відрізок (лівий чи правий) попадає X , ймовірність дорівнює:

$P=f(x(h))|dx(h)|+o(|dx(h)|)$ (змінюється лише аналітичний вираз для $o(|dx(h)|)$).

Примітка: Якщо попали в лівий відрізок - один вираз $0(|dx(h)|)$, якщо в правий – інший.

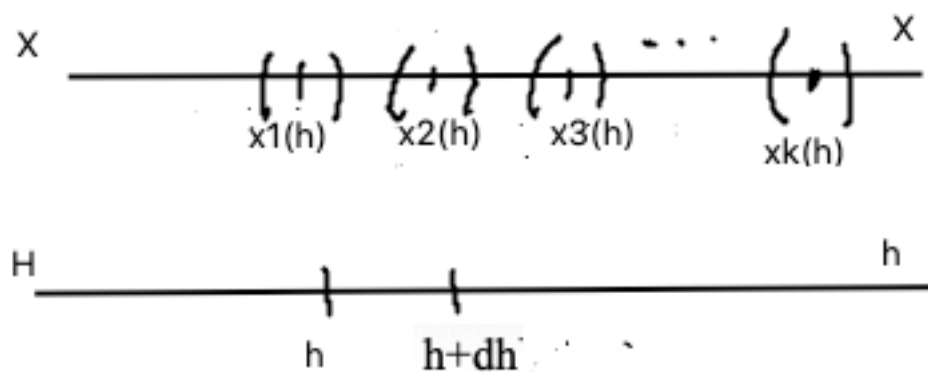
$$P=f(x(h))|dx(h)|+o(|dx(h)|)=g_H(h)dh + o(dh)$$

$$P=f(x(h))\frac{|dx(h)|}{dh} dh + o\left(\left|\frac{dx(h)}{dh}\right| dh\right)$$

$$\text{Якщо } \left|\frac{dx(h)}{dh}\right| < \infty, \text{ то } o\left(\left|\frac{dx(h)}{dh}\right| dh\right) = o(dh)$$

звідки, $g_H(h) = f(x(h))\frac{|dx(h)|}{dh}$.

У загальному випадку, коли у функції $h(x)$ - k -різних обернених функцій:



Нехай $h, h+dh$ – довільні, а $x_i(h)$ – різні.

У цьому випадку подія: H попала в $[h, h+dh]$ дорівнює об'єднанню k попарно несумісних подій: випадкова величина X попала у

$$[x_i(h)-|dx_i(h)|, x_i(h)], \text{ чи у } [x_i(h), x_i(h)+|dx_i(h)|], i = \overline{1, k}.$$

Звідки,

$$g_H(h) = \sum_{i=1}^k f(x_i(h)) \left| \frac{dx_i(h)}{dh} \right|.$$

Якщо існує $i \neq j$, $x_i(h) = x_j(h)$ (існують дві різні обернені функції з однаковими значеннями), то задача не розв'язується, тобто для даного h знайти функцію $g_H(h)$ неможливо.

ЛЕКЦІЯ 21

Нерівність Чебишова

$$P(Z \geq \tau) \leq \frac{MZ}{\tau}, \quad Z \geq 0, \quad \tau > 0. \quad (21.1)$$

Доведемо нерівність для неперервного випадку (для дискретного – довести самостійно).

$$f(z) = \begin{cases} 0, & z < 0 \\ f(z), & z \geq 0 \end{cases} \text{ – функція щільності}$$

$$MZ = \int_{-\infty}^{\infty} zf(z)dz = \int_0^{\infty} zf(z)dz = \int_0^{\tau} zf(z)dz + \int_{\tau}^{\infty} zf(z)dz \geq \int_{\tau}^{\infty} zf(z)dz \geq \tau \int_{\tau}^{\infty} f(z)dz = \tau P(Z \geq \tau).$$

Часткові випадки

Розглянемо випадкову величину $Z = |X - a|$, де X – деяка неперервна випадкова величина, a – деяке число. Тоді зауважимо, що $|X - a| \geq \varepsilon$ та $(X - a)^2 \geq \varepsilon^2$ є дві еквівалентні події, звідки $P(|X - a| \geq \varepsilon) = P((X - a)^2 \geq \varepsilon^2)$.

Застосуємо вже одержану нерівність (21.1), матимемо:

$$P(|X - a| \geq \varepsilon) = P((X - a)^2 \geq \varepsilon^2) \leq \frac{M(X - a)^2}{\varepsilon^2}. \quad (21.2)$$

Це другий варіант нерівності Чебишова.

У окремому випадку, коли $a = v = MX$, маємо

$$P(|X - v| \geq \varepsilon) \leq \frac{DX}{\varepsilon^2}. \quad (21.3)$$

Ця формула має лише теоретичне значення, не має практичного тому, що вірна для будь-якої випадкової величини. Отже, для конкретної – є завищеною.

Нехай X – деяка випадкова величина з $DX = 0$, тоді $P(|X - v| \geq \varepsilon) = 0$ для довільного $\varepsilon > 0$. Звідси виходить, що така випадкова величина з імовірністю одиниця є константою, яка дорівнює v .

Примітка. Перший випадок, коли ми отримали результат з ймовірністю 1.

*****Багатовимірні випадкові величини**

Розглянемо ймовірнісний простір (Ω, σ, P) .

n -вимірною випадковою величиною називається множина n одновимірних випадкових величин

$$\varphi_1(w), \varphi_2(w), \dots, \varphi_n(w).$$

$\varphi_i(w)$ – функції, вимірні відносно σ -алгебри

З означення n -вимірної випадкової величини випливає існування n -вимірної функції розподілу:

$$F(x_1, \dots, x_n) = P(\forall w: \varphi_1(w) < x_1, \varphi_2(w) < x_2, \dots, \varphi_n(w) < x_n).$$

Для довільних $x_1, x_2, \dots, x_n$, що належить R_n , маємо

$$A_i = \{ \omega: \varphi_i(\omega) < x_i \} \in \sigma, \quad i=1, \dots, n,$$

тобто A_i є подіями. Тому, з властивостей σ -алгебри виходить, що

$$A = \bigcap_{i=1}^n A_i = \bigcap_{i=1}^n \{ \omega: \varphi_i(\omega) < x_i \} \in \sigma,$$

тобто A є теж подією, для якої означена ймовірність $P(\cdot)$.

Властивості n -вимірної функції розподілу

1. Функція розподілу приймає значення в сегменті $[0; 1]$.
2. Якщо хоч один елемент функції розподілу дорівнює $-\infty$, то n -вимірна функція дорівнює нулю як ймовірність неможливої події.
3. Якщо всі елементи дорівнює $+\infty$, то n -вимірна функція дорівнює одиниці як ймовірність достовірної події.
4. n -вимірна функція розподілу – неспадна по будь-якій підмножині своїх аргументів.
5. n -вимірна функція розподілу – неперервна зліва по всім своїм аргументам.

Неперервність зліва будемо доводити при $n=2$.

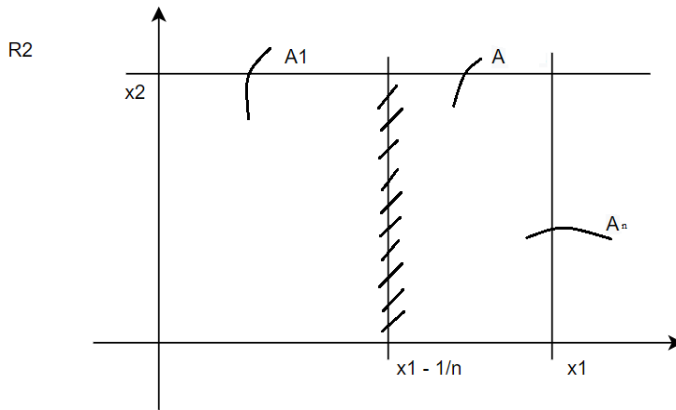
Логіка доведення

Спочатку доводимо рівність

$$F(x_1, x_2) = F(x_1 - 0, x_2) = F(x_1, x_2 - 0).$$

Потім наступну:

$$F(x_1, x_2) = F(x_1 - 0, x_2 - 0).$$



$$A = \{\forall \omega: \varphi_1(\omega) < x_1, \varphi_2(\omega) < x_2\}$$

$$A_n = \left\{ \forall \omega: x_1 - \frac{1}{n} \leq \varphi_2(\omega) < x_1 \right. \\ \left. \varphi_2(\omega) < x_2 \right\}$$

$$A_1 = \left\{ \forall \omega: \varphi_1(\omega) < x_1 - \frac{1}{n} \right. \\ \left. \varphi_2(\omega) < x_2 \right\}$$

В силу того, що $\varphi_1(\omega), \varphi_2(\omega)$ – скалярні

функції A_1, A_n – несумісні, а $A = A_1 \cup A_2$

$$P(A_n) = P(A) - P(A_1) = F(x_1, x_2) - F\left(x_1 - \frac{1}{n}, x_2\right),$$

де $F(x_1, x_2) = P(\forall \omega: \varphi_1(\omega) < x_1, \varphi_2(\omega) < x_2)$

Розглянемо наступний перетин:

$$\bigcap_{n=1}^{\infty} \left\{ \left[x_1 - \frac{1}{n}; x_1 \right) \oplus (-\infty; x_2) \right\} = \{\emptyset\}$$

$$\bigcap_{n=1}^{\infty} A_n = V - \text{неможлива подія} \Rightarrow P\left(\bigcap_{n=1}^{\infty} A_n\right) = 0$$

$A_{n+1} \subset A_n$ за побудовою.

Використовуємо теорему, яка буде доведена в посиленому законі великих чисел:

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

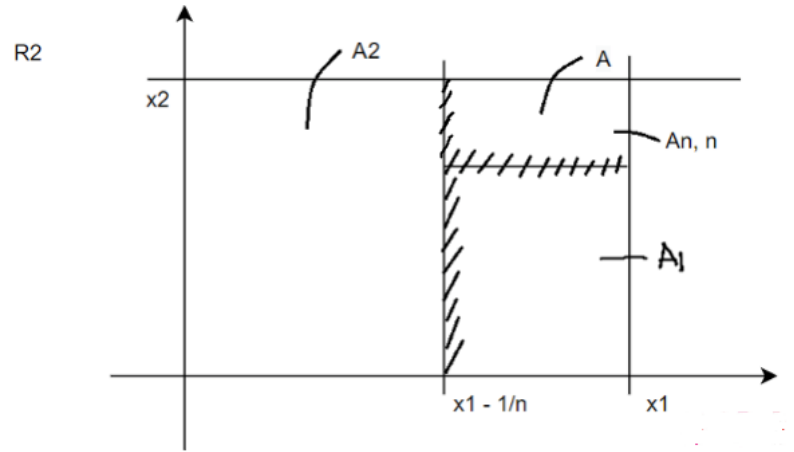
$$\lim_{n \rightarrow \infty} P(A_n) = 0 \Rightarrow \lim_{n \rightarrow \infty} \left(F(x_1, x_2) - F\left(x_1 - \frac{1}{n}, x_2\right) \right) = 0$$

$$F(x_1, x_2) = F(x_1 - 0, x_2).$$

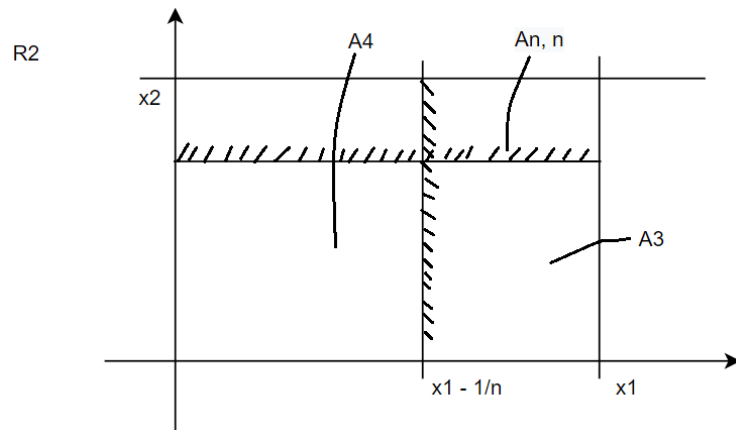
Таким чином, ми довели, що $F(x_1, x_2)$ – неперервна зліва по першому аргументу.

Аналогічно доводиться для x_2 .

Доводимо неперервність функції розподілу по двох аргументам:



Малюнок 22.1



Малюнок 22.2

$$A = \{\forall w: \varphi_1(w) < x_1, \varphi_2(w) < x_2\}$$

$$A_{n,n} = \left\{ \forall w: x_1 - \frac{1}{n} \leq \varphi_1(w) < x_1, x_2 - \frac{1}{n} \leq \varphi_2(w) < x_2 \right\}$$

$$A_1 = \left\{ \forall w: x_1 - \frac{1}{n} \leq \varphi_1(w) < x_1, \varphi_2(w) < x_2 - \frac{1}{n} \right\}$$

$$A_2 = \left\{ \forall w: \varphi_1(w) < x_1 - \frac{1}{n}, \varphi_2(w) < x_2 \right\}$$

$$A_3 = \left\{ \forall w: \varphi_1(w) < x_1, \varphi_2(w) < x_2 - \frac{1}{n} \right\}$$

$$A_4 = \left\{ \forall w: \varphi_1(w) < x_1 - \frac{1}{n}, \varphi_2(w) < x_2 - \frac{1}{n} \right\}$$

$A_1, A_2, A_{n,n}$ - попарно несумісні.

$$A = A_1 \cup A_2 \cup A_{n,n}$$

$$P(A_{n,n}) = P(A) - P(A_1) - P(A_2)$$

$$A_3 = A_4 \cup A_1$$

A_1, A_4 – несумісні.

$$P(A_1) = P(A_3) - P(A_4)$$

$$P(A_{n,n}) = P(A) - P(A_2) - P(A_3) + P(A_4)$$

$$P\left(\forall w: x_1 - \frac{1}{n} \leq \varphi_1(w) < x_1, x_2 - \frac{1}{n} \leq \varphi_2(w) < x_2\right) = F(x_1; x_2) - \\ - F\left(x_1 - \frac{1}{n}; x_2\right) - F\left(x_1; x_2 - \frac{1}{n}\right) + F\left(x_1 - \frac{1}{n}; x_2 - \frac{1}{n}\right)$$

Таким чином, ми фактично вивели наступну формулу:

$$P(\forall w: a_1 \leq \varphi_1(w) < b_1, a_2 \leq \varphi_2(w) < b_2) = F(b_1; b_2) - F(a_1; b_2) - F(b_1; a_2) + \\ F(a_1; a_2);$$

В нашому випадку:

$$a_1 = x_1 - \frac{1}{n} \quad b_1 = x_1$$

$$a_2 = x_2 - \frac{1}{n} \quad b_2 = x_2.$$

Розглянемо:

$$\bigcap_{n=1}^{\infty} \{[x_1 - \frac{1}{n}; x_1) \oplus [x_2 - \frac{1}{n}; x_2)\} = \{\emptyset\}$$

$$\bigcap_{n=1}^{\infty} A_{n,n} = V - \text{неможлива подія} \Rightarrow P\left(\bigcap_{n=1}^{\infty} A_{n,n}\right) = 0.$$

$$A_{n+1,n+1} \subset A_n$$

$$0 = \lim_{n \rightarrow \infty} P(A_{n,n}) = F(x_1; x_2) - F(x_1 - 0; x_2) - F(x_1; x_2 - 0) + F(x_1 - 0; x_2 - 0) \\ = 0.$$

Таким чином, ми довели, що n -вимірна функція розподілу неперервна зліва по всім аргументам $F(x_1, x_2) = F(x_1 - 0, x_2 - 0)$.

Так само доводиться, що існує:

$$P(\forall w: a_i \leq \varphi_1(w) < b_i), i = \overline{1, n}.$$

Ця ймовірність дорівнює сумі (різниці) значень n -вимірної функції розподілу по всім вершинам n -вимірного напівінтервалу, кількість яких - 2^n .

Як і в одновимірному випадку розглянемо множину K всіх різних підмножин арифметичного простору R^n .

$$K = \{B\}, B \subset R^n, \forall B \\ \{\forall w: \{\varphi_i(w), i = \overline{1, n} \in B\} \in \sigma\}.$$

Примітка: мається на увазі, що вектор $(\varphi_1(w), \varphi_2(w), \dots, \varphi_n(w))^T \in B$.

Нам відомо, що K містить сам арифметичний простір, а також всі можливі n -вимірні напівінтервали.

Аналогічно одновимірному випадку, доводимо $K \in \sigma$ – алгеброю, що містить алгебру всіх n -вимірних напівінтервалів, а значить містить мінімальну σ -алгебру $(\mathcal{B}_n)$ – борелівська алгебра на R^n .

$\mathcal{B}_n \subseteq K$ – це включення ми раніше довели для одновимірного випадку.

Задається скінчено-адитивна ймовірнісна міра на алгебрі пів-інтервалів і доводиться, що вона є злічено-адитивною на алгебрі напівінтервалів і за теоремою Каратеодорі, однозначно поширюється на $\mathcal{B}_n$.

Таким чином, якщо є $\varphi_1(w) \dots \varphi_n(w)$, то її n -вимірна функція розподілу дозволяє використовувати такий ймовірнісний простір:

$$(R_n, \mathcal{B}_n, P_{F(x_1, \dots, x_n)}).$$

Наслідком випробування над n – вимірною випадковою величиною є вектор з простору R_n .

Елементи $\mathcal{B}_n$ є всі можливі n -вимірні напівінтервали, інтервали, сегменти, вектори, всі можливі скінченні та нескінченні об'єднання і протилежні події, а також строго доводиться всі області, для яких існує n -вимірний об'єм.

ЛЕКЦІЯ 23

Означення. Абстрактна числова скалярна функція n -дійсних аргументів має характерні властивості n -вимірної функції розподілу, якщо:

- а) приймає значення з сегменту $[0, 1]$;
- б) неспадна по будь-якій підмножині своїх аргументів;

- в) неперервна зліва по всім своїм аргументам;
- г) якщо хоч один аргумент дорівнює $-\infty$, то функція дорівнює 0;
- д) якщо всі аргументи дорівнюють $+\infty$, то функція дорівнює 1.

Тому, повністю аналогічно тому, як було показано для числової скалярної функції дійсного скалярного аргументу, що має характеристичні властивості одновимірної функції розподілу, має місце наступний результат.

Функція $F(x_1, \dots, x_n)$ дозволяє побудувати абстрактний ймовірнісний простір $(\mathbb{R}_n, \mathcal{B}_n, P_{F(x_1, \dots, x_n)})$, що задає абстрактну n -вимірну випадкову величину, для якої ця функція є функцією розподілу (обернена теорема Колмогорова).

Означення. Числова скалярна функція $g(x_1, \dots, x_n)$ називається борелівською, якщо для будь-якого дійсного числа y множина всіх дійсних x (векторів R_n) для яких виконується

$$\forall x \in R_n \quad g(x_1, \dots, x_n) < y = B \in \mathcal{B}_n \quad (\text{є елементом борелівської алгебри}).$$

Тоді, як і в одновимірному випадку, доводиться, що якщо $(X_1, \dots, X_n)$ – n -вимірна випадкова величина, то $g(X_1, \dots, X_n)$ є одновимірною випадковою величиною.

Примітка. Борелівські функції – всі аналітичні, неперервні, кусково-неперервні з обмеженою кількістю стрибків 1-го роду.

Двовимірні дискретні випадкові величини

Означення. Двовимірна дискретна випадкова величина – це сукупність двох одновимірних випадкових величин X та Y

$$X \sim \begin{pmatrix} x_1 & \dots & x_s \\ P_x(x_1) & \dots & P_x(x_s) \end{pmatrix}$$

$$Y \sim \begin{pmatrix} y_1 & \dots & y_m \\ P_y(y_1) & \dots & P_y(y_m) \end{pmatrix}$$

Примітка 1. Для спрощення розглянемо випадок, коли s та m – обмежені натуральні числа.

Примітка 2. Ми вважаємо, що для X і Y має місце ймовірнісний простір $(\mathbb{R}_2, \mathcal{B}_2, P_{F(x_1, x_2)})$. Тоді множина можливих значень двовимірної дискретної випадкової величини – це множина всіх векторів $\{x_i, y_j\}_{j=\overline{1, m}}$ на площині.

Гарантовано існує $P(x_i, y_j)$ для будь-якої пари x_i і y_j , тому що будь-який вектор площини є елементом борелівської алгебри $\mathcal{B}_2$ і має ймовірність настання $P_{F(x_1, x_2)}$.

Якщо пара чисел не належить цій множині, то вона також має ймовірність настання, що дорівнює 0 (неможлива подія), також будь-яка інтегровна область $\mathbb{R}_2$, що не містить жодної точки $\{x_i, y_j\}_{j=\overline{1, m}}$ має ймовірність настання, яка дорівнює нулю.

Подія A – внаслідок випробування над двовимірною випадковою величиною (X, Y) випадкова величина X прийняла значення x_i ($X = x_i$)

$$A = \bigcup_{j=1}^m \{x_i, y_j\}$$

$$P(A) = \sum_{j=1}^m P(x_i, y_j) = P_x(x_i)$$

Подія B – внаслідок випробування над двовимірною випадковою величиною (X, Y) випадкова величина Y прийняла значення y_j ($Y = y_j$).

$$B = \bigcup_{i=1}^s \{x_i, y_j\}$$

$$P(B) = \sum_{i=1}^s P(x_i, y_j) = P_y(y_j)$$

$$A \cap B = \{x_i, y_j\}$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(x_i, y_j)}{P_x(x_i)}$$

$$P(Y = y_j / X = x_i) = P(y_j / x_i) = \frac{P(x_i, y_j)}{P_x(x_i)}$$

$$\sum_{j=1}^m P(y_j / x_i) = \frac{\sum_{j=1}^m P(x_i, y_j)}{P_x(x_i)} = 1$$

$$P(X = x_i / Y = y_j) = P(x_i / y_j) = \frac{P(x_i, y_j)}{P_y(y_j)}$$

$$\sum_{i=1}^m P(x_i / y_j) = 1$$

Важливими поняттями в теорії ймовірностей є умовне математичне сподівання та умовна дисперсія.

Означення. Умовним математичним сподіванням випадкової величини Y , за умови, що випадкова величина X одержує значення x_i , називається сума:

$$M(Y / X = x_i) = M(Y / x_i) = \bar{y}(x_i) = \sum_{j=1}^m y_j P(y_j / x_i)$$

$$MY = \sum_{j=1}^m y_j P_y(y_j)$$

$$M(X / Y = y_j) = M(X / y_j) = \bar{x}(y_j) = \sum_{i=1}^s x_i P(x_i / y_j)$$

Означення : Умовною дисперсією випадкової величини Y , за умови, що випадкова величина X приймає значення x_i , називається сума

$$D(Y / x_i) = \sigma_{Y/x_i}^2 = \sum_{j=1}^m (y_j - \bar{y}(x_i))^2 P(y_j / x_i).$$

$$D(X / y_j) = \sigma_{X/y_j}^2 = \sum_{i=1}^s (x_i - \bar{x}(y_j))^2 P(x_i / y_j)$$

$$DY = \sum_{j=1}^m (y_j - MY)^2 P_y(y_j).$$

Умовне математичне сподівання відрізняється від безумовного тим, що в кожен добуток суми входить умовна ймовірність настання числа, а не безумовна.

Умовна дисперсія відрізняється від безумовної тим, що у вираз дисперсії замість безумовного математичного сподівання пишеться умовне, замість безумовної ймовірності – умовна.

Зміст умовного математичного сподівання.

Умовне математичне сподівання – це число на числовій осі, відносно якого групуються конкретні результати випробувань над однією випадковою величиною, якщо в цьому випробуванні друга випадкова величина приймає одне й теж саме фіксоване значення (за умови, що кількість таких випробувань достатньо велика).

Зміст умовної дисперсії.

Умовна дисперсія – це якісна міра ступеня концентрації результатів конкретних випробувань над однією випадковою величиною відносно її умовного математичного сподівання, якщо в кожному випробуванні друга випадкова величина приймає одне й теж саме фіксоване значення.

Чим менша умовна дисперсія, тим тісніша ступінь концентрації.

ЛЕКЦІЯ 24

Двовимірні неперервні випадкові величини

Двовимірна неперервна випадкова величина – це сукупність двох одновимірних неперервних випадкових величин.

$$\begin{array}{cc} X & Y \\ F_X(x) & F_Y(y) \\ f_X(x) & f_Y(y). \end{array}$$

Відомо, що для двовимірної неперервної випадкової величини існує двовимірна функція розподілу

$$F(x, y) = P(X < x, Y < y).$$

Покажемо:

$$P(X < x, Y < y) = P\left(X \leq x, Y \leq y\right) -$$

очевидно, що розглядаються 4 варіанти.

Дана рівність є узагальненням результату для одновимірного варіанту

$$P(X < x) = P(X \leq x).$$

Доведення для кожного варіанту доведення однотипне, тож розглянемо лише один варіант.

Покажемо, що:

$$P(X \leq x, Y \leq y) = P(X < x, Y \leq y).$$

Подія $(X \leq x, Y \leq y)$ складається з усіх пар чисел u та v , які задовольняють нерівності: $u \leq x, v \leq y$, а $(X < x, Y \leq y)$ складається з усіх пар чисел u та v , які задовольняють нерівності: $u < x, v \leq y$.

Ймовірнісний простір задається трійкою $(R_2, B_2, P_{F(x,y)})$.

Подія $(X \leq x, Y \leq y)$ представляється у вигляді об'єднання наступних подій:

$$(X \leq x, Y \leq y) = (X < x, Y \leq y) \cup (X = x, Y \leq y).$$

Якщо $A \subset B$, то $P(A) \leq P(B)$.

Подія $(X = x, Y \leq y) \subset (X = x, Y < \infty)$,

але $P(X = x, Y < \infty) = P(X = x) = 0$, за властивістю неперервної випадкової величини.

Таким чином, $P(X \leq x, Y \leq y) = P(X < x, Y \leq y)$ (застосували третю аксіому Колмогорова).

Інші три варіанти доводяться аналогічно.

Наслідок 1. Двовимірна функція розподілу неперервна.

Наслідок 2. Ймовірність настання для усіх чотирьох варіантів

$$P\left(a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2\right) = const.$$

Доведення базується на методології доведення рівності

$$P(X \leq x, Y \leq y) = P(X < x, Y \leq y).$$

Двовимірна випадкова величина називається **неперервною**, якщо її двовимірна функція розподілу є неперервною.

Двовимірною випадковою величиною називається **абсолютно неперервною**, якщо існує така числова скалярна функція, що належить до класу неперервних, кусково-неперервних функцій з обмеженою кількістю стрибків I-го роду і задовольняє наступній інтегральній рівності.

$$F(x, y) = \int_{-\infty}^x \left[\int_{-\infty}^y f(u, v) dv \right] du = \iint_{(\infty; x) \oplus (-\infty; y)} f(u, v) du dv. \quad (24.1)$$

Примітка:

$$F(x) = \int_{-\infty}^x f_x(u) du.$$

Примітка. Коли ми будемо писати двократний (n-кратний) інтеграл, диференціали будемо розташовувати в наступному порядку: спочатку зовнішній, потім внутрішній.

Властивості двовимірної неперервної функції щільності

Якщо для двовимірної випадкової неперервної величини існує функція щільності з класу неперервних чи кусково-неперервних функцій з обмеженою кількістю стрибків і задовольняє інтегральній рівності (24.1), то вона **абсолютно неперервна**.

1. Функція щільності невід'ємна $f(x, y) \geq 0$, бо функція розподілу неспадна по всім своїм аргументам.
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = 1$, як ймовірність достовірної події.
3. Нехай в області G вона неперервна, тоді для будь-якої точки з G інтегральна рівність (24.1) перетворюється в наступну:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y},$$

– друга змішана часткова похідна по аргументам x, y.

Примітка: Для одновимірного випадку

$$f(x) = \frac{dF(x)}{dx}, \quad F(x) = \int_{-\infty}^x f(u) du.$$

За означенням:

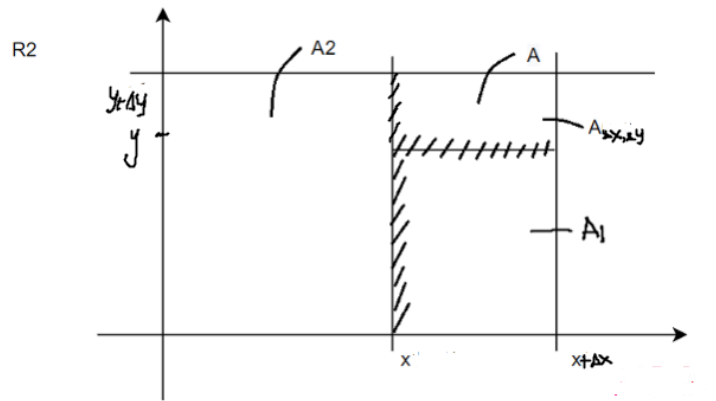
$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) - F(x, y + \Delta y) + F(x, y)}{\Delta x \Delta y}.$$

Виведемо наступну формулу:

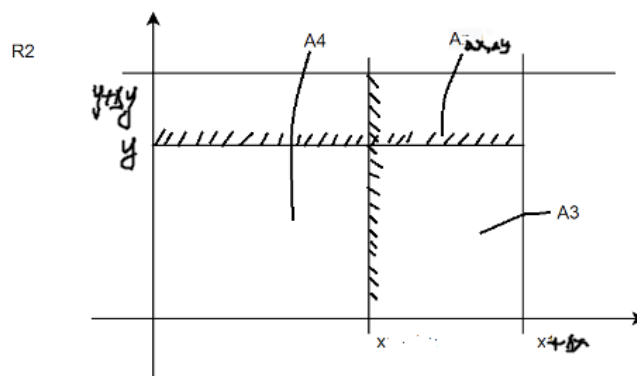
$$P\left(\begin{matrix} x \leq X \leq x + \Delta x \\ y \leq Y \leq y + \Delta y \end{matrix}\right) = F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) - F(x, y + \Delta y) + F(x, y).$$

Примітка. Як показано вище

$$P\left(\begin{matrix} x \leq X \leq x + \Delta x \\ y \leq Y \leq y + \Delta y \end{matrix}\right) = P\left(\begin{matrix} x \leq X < x + \Delta x \\ y \leq Y < y + \Delta y \end{matrix}\right)$$



Малюнок 24.1



Малюнок 24.2

$$A = \{X < x + \Delta x, Y < y + \Delta y\}$$

$$A_{\Delta x, \Delta y} = \{x \leq X < x + \Delta x, y \leq Y < y + \Delta y\}$$

$$A_1 = \{x \leq X < x + \Delta x, Y < y\}$$

$$A_2 = \{X < x, Y < y + \Delta y\}$$

$$A_3 = \{X < x + \Delta x, Y < y\}$$

$$A_4 = \{X < x, Y < y\}$$

$A_1, A_2, A_{\Delta x, \Delta y}$ - попарно несумісні.

$$A = A_1 \cup A_2 \cup A_{\Delta x, \Delta y}$$

$$P(A_{\Delta x, \Delta y}) = P(A) - P(A_1) - P(A_2)$$

$$A_3 = A_4 \cup A_1$$

A_1, A_4 – несумісні.

$$P(A_1) = P(A_3) - P(A_4)$$

$$P(A_{\Delta x, \Delta y}) = P(A) - P(A_2) - P(A_3) + P(A_4)$$

$$P(x \leq X < x + \Delta x, y \leq Y < y + \Delta y) = F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y)$$

Примітка. Таким чином, ми фактично вивели наступну формулу:

$$P(a_1 \leq X < b_1, a_2 \leq Y < b_2) = F(b_1; b_2) - F(a_1; b_2) - F(b_1; a_2) + F(a_1; a_2);$$

В нашому випадку:

$$a_1 = x \quad b_1 = x + \Delta x$$

$$a_2 = y \quad b_2 = y + \Delta y.$$

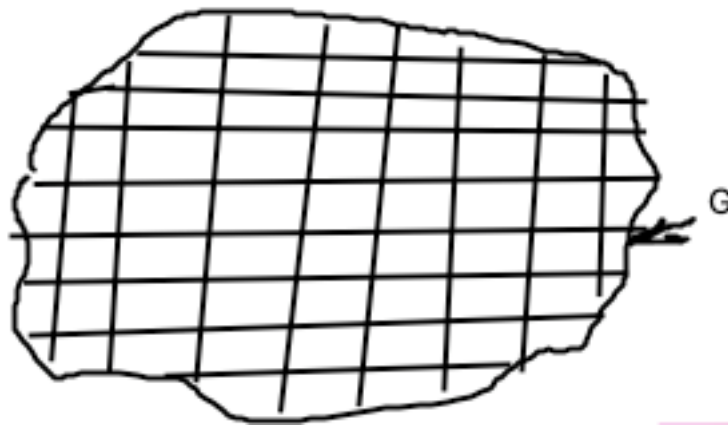
Таким чином

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x \geq 0 \\ \Delta y \rightarrow 0 \\ \Delta y \geq 0}} \frac{P\left(\begin{matrix} x \leq X \leq x + \Delta x \\ y \leq Y \leq y + \Delta y \end{matrix}\right)}{\Delta x \Delta y}$$

$$4. P\left(\begin{matrix} x \leq X \leq x + \Delta x \\ y \leq Y \leq y + \Delta y \end{matrix}\right) = f(x, y) \Delta x \Delta y + O(\Delta x \Delta y)$$

як аналог $P(x \leq X \leq x + \Delta x) = f_x(x) \Delta x + O(\Delta x)$

$$5. P(XY \in G) = \iint_G f(x, y) dx dy \quad (24.2)$$



Обґрунтування: область G представляємо об'єднанням двовимірних напівінтервалів, що не перетинаються.

Ймовірність попадання в кожен напівінтервал дорівнює:

$$f(x, y)\Delta x\Delta y + O(\Delta x\Delta y),$$

x, y належить цьому напівінтервалу.

Наближено $P(X, Y \in G)$ дорівнює сумі ймовірностей попадання в усі напівінтервали.

Точну формулу отримаємо, спрямувавши сторони прямокутників $\Delta x, \Delta y$ до 0.

Примітка: Означення функції щільності двовимірної випадкової величини є частковим випадком (24.2), де $G = (\infty; x) \oplus (-\infty; y)$.

ЛЕКЦІЯ 25

Маємо двовимірну випадкову величину (X, Y) та функцію щільності $f(x, y)$.

Необхідно знайти $f_X(x), f_Y(y)$.

Примітка. З точки зору практики – задача безглузда, але виведена формула буде використана в різних теоретичних випадках.

Розв'яжемо задачу:

$$P(X \leq x) = \int_{-\infty}^x f_X(u) du$$

$$P(X \leq x, Y < \infty) = \int_{-\infty}^x \left[\int_{-\infty}^{\infty} f(u, v) dv \right] du = P(X \leq x) = \int_{-\infty}^x f_X(u) du$$

$$(P(X, Y \in G) = \iint_G f(x, y) dx dy)$$

$$f_X(u) = \int_{-\infty}^{\infty} f(u, v) dv$$

Ми прирівняли підінтегральні вирази тільки тому, що ця інтегральна рівність має місце $\forall x \in \mathbb{R}$. Це означає, що в нас не одна інтегральна рівність, а нескінченно-незлічена множина інтегральних рівностей (для будь-якого дійсного x). Тоді, строго можна довести що, якщо підінтегральні функції неперервні чи кусково-неперервні з обмеженою кількістю стрибків I-го роду, це можливо тільки тоді, коли вони рівні.

Має місце симетрична формула:

$$f_y(v) = \int_{-\infty}^{\infty} f(u, v) du.$$

Таким чином, отримали наступний результат: щоб по двовимірній функції щільності знайти одновимірну функцію щільності однієї випадкової величини, треба двовимірну функцію щільності проінтегрувати по нескінченним границям по змінній, що відповідає другій випадковій величині.

Примітка. З означення функції щільності випливає, що якщо задана двовимірною випадкова величина (X, Y) , то перший аргумент відповідає X , а другий – Y .

Умовна функція щільності

$$P(y \leq Y \leq y + \Delta y) = f_Y(y) + o(\Delta y)$$

$$P(y \leq Y \leq y + \Delta y / X = x) = f(y/x)\Delta y + o(\Delta y)$$

$$f(y/x) = \frac{f(x, y)}{f_X(x)}$$

Остання рівність визначає умовну функцію щільності випадкової величини Y за умови, що випадкова величина X одержує значення x .

Примітка. Довести цю формулу неможливо, як і формулу умовної ймовірності. Але її обґрунтування зовсім інше, ніж у формули умовної ймовірності.

Двовимірною випадковою величиною задано на $(\mathbb{R}_2, \mathcal{B}_2, P_{F(x,y)})$, розглянемо події A і B :

$$A = (x \leq X \leq x + \Delta x, Y < \infty)$$

$$B = (X < \infty, y \leq Y \leq y + \Delta y)$$

$$A \cdot B = \{x \leq X \leq x + \Delta x, y \leq Y \leq y + \Delta y\}$$

Ймовірності цих подій дорівнюють:

$$P(A) = f_X(x)\Delta x + o(\Delta x),$$

$$P(B) = f_Y(y)\Delta y + o(\Delta y),$$

$$P(AB) = f(x, y)\Delta x\Delta y + o(\Delta x \cdot \Delta y).$$

Умовна ймовірність $P(B / A)$ дорівнює

$$P(B/A) = P(y \leq Y \leq y + \Delta y / x \leq X \leq x + \Delta x) = \frac{P(A \cdot B)}{P(A)} =$$

$$= \frac{f(x, y) \Delta x \Delta y + o(\Delta x \cdot \Delta y)}{f_x(x) \Delta x + o(\Delta x)} = \frac{f(x, y)}{f_x(x)} \Delta y + o(\Delta y).$$

Доведення:

$$\left(\frac{f(x, y)}{f_x(x)} \Delta y + o(\Delta y) \right) (f_x(x) \Delta x + o(\Delta x)) =$$

$$= f(x, y) \Delta x \Delta y + o(\Delta x \cdot \Delta y).$$

Примітка. Довели для випадку $x \leq X \leq x + \Delta x$, а не для випадку $X = x$.

Таким чином, умовні функції щільності мають вигляд

$$f_{Y/x}(y/x) = \frac{f(x, y)}{f_x(x)}$$

$$f_{X/y}(x/y) = \frac{f(x, y)}{f_y(y)}$$

Означення. Умовним математичним сподіванням випадкової величини Y за умови, що X одержує значення x , називають величину

$$M(Y/X = x) = \bar{y}(x) = \int_{-\infty}^{\infty} y f_{Y/x}(y/x) dy.$$

Умовне математичне сподівання X визначається рівністю

$$M(X/Y = y) = \bar{x}(y) = \int_{-\infty}^{\infty} x f_{X/y}(x/y) dx.$$

Умовною дисперсією величини Y за умови, що X одержує значення x , називається величина

$$\sigma_{Y/x}^2 = \int_{-\infty}^{\infty} (y - \bar{y}(x))^2 f_{Y/x}(y/x) dy,$$

$$\sigma_{X/y}^2 = \int_{-\infty}^{\infty} (x - \bar{x}(y))^2 f_{X/y}(x/y) dx.$$

Умовне математичне сподівання і умовна дисперсія для неперервної випадкової величини відрізняються від безумовних тим, що у виразі для умовного математичного сподівання замість безумовної функції щільності ставиться умовна. А у формулі для умовної дисперсії замість безумовних математичного сподівання і функції щільності – умовне математичне сподівання і умовна функція щільності.

Примітка.

$\bar{y}(x)$ і $\bar{x}(y)$ називають лініями регресії.

Зміст умовного математичного сподівання і умовної дисперсії такий самий, як і у випадку дискретних випадкових величин.

ЛЕКЦІЯ 26

Двовимірні незалежні дискретні випадкові величини

Маємо XY – двовимірну дискретну випадкову величину.

Дискретні випадкові величини X , Y називаються **незалежними** в двовимірній дискретній випадковій величині XY , якщо

$$\forall_{i,j} P(x_i, y_j) = P_x(x_i)P_y(y_j), \text{ якщо } P_x(x_i) > 0, P_y(y_j) > 0, \text{ то}$$

$$P(y_j/x_i) = P_y(y_j), P(x_i/y_j) = P_x(x_i).$$

$$\Rightarrow M(Y/x_i) = MY, M(X/y_j) = MX,$$

оскільки

$$(P(y_j/x_i) = \frac{P(x_i y_j)}{P_x(x_i)}, P(x_i/y_j) = \frac{P(x_i y_j)}{P_y(y_j)}).$$

Нехай є два простих випробування, що не мають спільних випадкових факторів:

I. II.

X Y.

Тоді $\forall_{i,j} P(x_i, y_j) = P_x(x_i)P_y(y_j)$ отримуємо, використовуючи загальну теорію (композицію двох незалежних випробувань).

Дана формула читається наступним чином: ймовірність того, що в першому випробуванні випадкова величина X прийме значення x_i , а в другому випробуванні випадкова величина Y – значення y_j , дорівнює добутку ймовірностей їх настання.

Примітка: Якщо в першому випробуванні настає подія ($X \in A$), а в другому ($Y \in B$), то $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$.

Доведення випливає з загальної структури незалежних подій, породжених двома випробуваннями, що не мають спільних факторів.

Двовимірні незалежні неперервні випадкові величини

Маємо XY – двовимірна неперервна випадкова величина.

Неперервні випадкові величини X, Y називаються незалежними, якщо

$$F(x, y) = F_x(x) \cdot F_y(y).$$

Наслідок: Якщо в точці (x, y) двовимірна функція щільності неперервна, то еквівалентним означенням є:

Неперервні випадкові величини X, Y називаються незалежними, якщо

$$f(x, y) = f_x(x) \cdot f_y(y).$$

Доведення: Якщо в точці (x, y) , $F(x, y)$ – неперервна, то

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \cdot \partial y} = \frac{\partial^2 (F_x(x) \cdot F_y(y))}{\partial x \cdot \partial y} = f_x(x) \cdot f_y(y).$$

Легко переконатися, що умовна функція щільності дорівнює безумовній, умовне математичне сподівання дорівнює безумовному, умовна дисперсія дорівнює безумовній дисперсії.

Нехай є два простих випробування, що не мають спільних випадкових факторів:

I.	II.
X	Y .

Якщо в першому випробуванні настає подія ($X \in A$), а в другому ($Y \in B$)., то $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$ - отримуємо використовуючи загальну теорію (композицію двох незалежних випробувань).

Читається наступним чином: ймовірність того, що в першому випробуванні випадкова величина X прийме значення з A , а у другому випробуванні випадкова величина Y – значення з B , дорівнює добутку ймовірностей їх настання.

Розглянемо частковий випадок:

$$P(X < x, Y < y) = P_x(X < x) \cdot P_y(Y < y) = F_x(x) \cdot F_y(y),$$

$$P(X < x, Y < y) = F(x, y).$$

Формально X і Y незалежні, тому: $F(x, y) = F_x(x) \cdot F_y(y)$. Але в цьому випадку двовимірна функція розподілу $F(x, y)$ – числова скалярна функція двох дійсних аргументів, яка при фіксованих значеннях аргументів дорівнює $P(X < x, Y < y)$, що читається так: це ймовірність того, що в першому випробуванні $X < x$, а в другому випробуванні $Y < y$.

Багатовимірні дискретні випадкові величини

Нехай n – вимірна випадкова величина задається n одномірними випадковими величинами $X_1 \dots X_n$.

$$X_i \sim \begin{pmatrix} x_1^i, & \dots & x_{s_i}^i \dots \end{pmatrix}$$

$$P_{x_i}(x_1^i) \dots P_{x_i}(x_{s_i}^i)$$

$$i = \overline{1, n}$$

Ймовірнісний простір $(R_n, \mathcal{B}_n, P_{F(x_1 \dots x_n)})$.

Можливі значення – множина векторів $\{x_{j_1}^1, x_{j_2}^2 \dots x_{j_n}^n\}$

$$\forall j_i \in (1, \dots s_i)$$

$$i = \overline{1, n}$$

Оскільки всі вектори є елементами $\mathcal{B}_n$, для будь якого вектора існує ймовірність настання $P(x_{j_1}^1, x_{j_2}^2 \dots x_{j_n}^n)$.

Випадкові величини X_i – незалежні, якщо

$$\forall P\{x_{j_1}^1, x_{j_2}^2 \dots x_{j_n}^n\} = \prod_{i=1}^n P_{x_i}(x_{j_i}^i)$$

Нехай маємо n простих випробувань, будь які 2 з них не мають спільних випадкових факторів. Кожне з них породжує X_i , тоді:

$\forall P\{x_{j_1}^1, x_{j_2}^2 \dots x_{j_n}^n\} = \prod_{i=1}^n P_{x_i}(x_{j_i}^i)$ – це ймовірність ймовірність того, що в i – му випробуванні X_i , прийме значення $x_{j_i}^i$, $i = \overline{1, n}$ дорівнює добутку ймовірностей.

п-вимірна неперервна випадкова величина

Нехай n – вимірна випадкова величина задається n одномірними випадковими величинами $X_1 \dots X_n$, для яких існує функція розподілу $F_{x_i}(x_i)$. Задається трійка $(R_n, \mathcal{B}_n, P_{F(x_1 \dots x_n)})$.

Існує n вимірна функція розподілу $F(x_1 \dots x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$

Наслідок:

$$P(x_i \leq X_i \leq x_i + \Delta x_i, i = \overline{1, n}) = \text{const.}$$

Таким чином n – вимірна функція розподілу – неперервна.

Примітка: замість лівої границі інтервалу можна ставити $-\infty$.

Означення: n -вимірна неперервна випадкова величина називається **абсолютно неперервною**, якщо для неї існує $f(x_1 \dots x_n)$ з класу неперервних функцій чи кусково-неперервних функцій з обмеженою кількістю стрибків першого роду:

$$F(x_1 \dots x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(u_1 \dots u_n) du_n \dots du_1.$$

$f(u_1 \dots u_n)$ – n -вимірна функція щільності.

Якщо замість $x_1 \dots x_n$ взяти $+\infty$, то

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(u_1 \dots u_n) du_n \dots du_1 = 1.$$

Якщо в точці $(x_1 \dots x_n)$ функція щільності є неперервною, то аналогом двовимірного випадку є

$$f(x_1 \dots x_n) = \frac{\partial^n F(x_1 \dots x_n)}{\partial x_1 \dots \partial x_n}$$

З цієї формули випливає:

$$= \lim_{\substack{\Delta x_i \rightarrow 0, \Delta x_i > 0, \\ i = \overline{1, n}}} \frac{\sum \pm F(\dots)}{\Delta x_1 \dots \Delta x_n} = \lim_{\substack{\Delta x_i \rightarrow 0, \Delta x_i > 0, \\ i = \overline{1, n}}} \frac{P(x_i \leq X_i \leq x_i + \Delta x_i, i = \overline{1, n})}{\Delta x_1 \dots \Delta x_n},$$

$$P(x_i \leq X_i \leq x_i + \Delta x_i) = f(x_1 \dots x_n) \prod_{i=1}^n \Delta x_i + o\left(\prod_{i=1}^n \Delta x_i\right).$$

$$P(X_1 \dots X_n \in G) = \int \dots \int_G f(x_1 \dots x_n) dx_1 \dots dx_n, G \in R_n.$$

ЛЕКЦІЯ 27

Задача: Маємо n -вимірну функцію щільності $f(x_1, \dots, x_n)$ n -вимірної неперервної випадкової величини $(X_1, \dots, X_n)$.

Розглянемо довільну k -вимірну величину $(X_{i_1}, \dots, X_{i_k})$, $k < n$, де випадкові величини X_{i_j} , $j = \overline{1, k}$ взяті з множини $(X_1, \dots, X_n)$.

Знайти k -вимірну функцію щільності $f(x_{i_1}, \dots, x_{i_k})$.

Розв'язок.

Знайдемо $f(x_{i_1}, \dots, x_{i_k})$.

$$P(X_{i_1} < x_{i_1}, \dots, X_{i_k} < x_{i_k}) = \int_{-\infty}^{x_{i_1}} \dots \int_{-\infty}^{x_{i_k}} f(u_{i_1}, \dots, u_{i_k}) du_{i_k} \dots du_{i_1}.$$

Імовірність цієї події дорівнює імовірності наступної події

$$P(X_{i_1} < x_{i_1}, \dots, X_{i_k} < x_{i_k}, X_{j_1} < \infty, \dots, X_{j_{n-k}} < \infty).$$

Тоді відповідний n -вимірний інтеграл дорівнює n -кратному:

$$\int_{-\infty}^{x_{i_1}} \dots \int_{-\infty}^{x_{i_k}} \left[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, u_n) du_{j_{n-k}} \dots du_{j_1} \right] du_{i_k} \dots du_{i_1}.$$

Примітка. Порядок в n -кратному інтегралі звичайних інтегралів змінено, спочатку йдуть звичайні інтеграли по змінним $u_{i_k}, \dots, u_{i_1}$, а потім $u_{j_{n-k}}, \dots, u_{j_1}$.

З рівності n -кратного і k -кратного інтегралів отримаємо результат:

$$f(x_{i_1}, \dots, x_{i_k}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, u_n) du_{j_{n-k}} \dots du_{j_1}$$

Примітка. Підінтегральний вираз k -кратного інтегралу і вираз в квадратних дужках в n -кратному інтегралі рівні, так як наша інтегральна рівність має місце для будь-яких дійсних значень $x_{i_1}, \dots, x_{i_k}$.

n -вимірні неперервні незалежні випадкові величини

Означення. Неперервні випадкові величини $X_1, X_2, \dots, X_n$ — незалежні, якщо їх n -вимірна функція розподілу дорівнює добутку їх одновимірних функцій розподілу

$$F(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i).$$

Якщо n -вимірний функція щільності – неперервна, то

$$f(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i).$$

Нехай маємо n простих випробувань, кожне з яких породжує неперервну випадкову величину $X_i, i = \overline{1, n}$, у якої є своя функція розподілу і функція щільності $F_{X_i}(x_i), f_{X_i}(x_i)$. Будь-які два прості випробування не мають спільних випадкових факторів. Тоді має місце наступна формула:

$$F(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i).$$

значення n -вимірної функції розподілу дорівнює ймовірності того, що в першому випробуванні $X_1 \leq x_1$, в другому – $X_2 \leq x_2$, ..., в n -у випробуванні $X_n \leq x_n$.

Математичне сподівання від числової скалярної функції випадкових аргументів

1) Дискретний випадок, $n = 2$.

Нехай X_1, X_2 – двовимірний дискретний випадковий величина

$$X_i \sim \begin{pmatrix} x_{j_1}^i & \dots & x_{j_s}^i \\ P_i(x_{j_1}^i) & \dots & P_i(x_{j_s}^i) \end{pmatrix}$$

Нехай функція $\Psi(x_1, x_2)$ – неперервна чи кусково-неперервна з обмеженою кількістю стрибків I-го роду.

Примітка. Треба однозначно задати значення функції при аргументах, в яких вона має стрибок I-го роду.

$\Psi(X_1, X_2)$ – числова скалярна функція двох дискретних випадкових аргументів, її реалізація в випробуванні є число $\Psi(x_{j_1}^1, x_{j_2}^2)$ (якщо X_1 прийняло значення $x_{j_1}^1$, а X_2 – $x_{j_2}^2$).

Задача. Знайти математичне сподівання цієї одновимірної випадкової величини $M\Psi(X_1, X_2)$. Побудуємо таблицю:

$$\left\{ \begin{array}{l} \Psi(x_{j_1}^1, x_{j_2}^2), j_1 = \overline{1, s_1} \\ P(x_{j_1}^1, x_{j_2}^2), j_2 = \overline{1, s_2} \end{array} \right\}$$

Примітка. Це не є таблицею для одновимірної випадкової величини $\Psi(X_1, X_2)$, так як в першому рядку значення можуть повторюватись.

Аналогічно одновимірному випадку отримуємо формулу:

$$M\Psi(X_1, X_2) = \sum_{j_1=1}^{s_1} \sum_{j_2=1}^{s_2} \Psi(x_{j_1}^1, x_{j_2}^2) P(x_{j_1}^1, x_{j_2}^2)$$

(аналог формули $\sum_{i=1}^s \Psi(x_i) P(x_i)$).

Неперервний випадок, $n = 2$.

X_1, X_2 – двовимірна неперервна випадкова величина, що має функцію щільності $f(x_1, x_2)$.

$$M\Psi(X_1, X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi(x_1, x_2) f(x_1, x_2) dx_1 dx_2.$$

Примітка. Узагальнення формули

$$M\Psi(X) = \int_{-\infty}^{\infty} \Psi(x) f(x) dx.$$

Для спрощення обґрунтування будемо вважати, що $\Psi(X_1, X_2)$ – неперервна функція і двовимірна функція щільності також неперервна. Якщо це не так, то площину R_2 , треба представити як об'єднання областей G_i , в кожній з яких ці функції неперервні.

Площину розбиваємо на велику кількість прямокутників Δ_{ij} , $i = \overline{-\infty, \infty}$; $j = \overline{-\infty, \infty}$. Якщо візьмемо один прямокутник, то Δx_i^1 і Δx_j^2 – сторони цього прямокутника.

X_1, X_2 апроксимуємо двовимірною дискретною X_1^*, X_2^* . Якщо у випробуванні X_1, X_2 прийняли значення, що належить Δ_{ij} , то в цьому випробуванні X_1^* прийняло

значення x_i^1 , а X_2^* прийняло значення x_j^2 . Зрозуміло, чим менше сторони прямокутника Δx_i^1 і Δx_j^2 , тим точніше випадкова величина $\Psi(X_1, X_2)$ апроксимується дискретною випадковою величиною $\Psi(X_1^*, X_2^*)$.

Для дискретної випадкової величини ймовірність прийняти значення $\Psi(x_i^1, x_j^2)$ дорівнює ймовірності потрапляння внаслідок випробування двовимірної випадкової величини X_1, X_2 у прямокутник Δ_{ij} і дорівнює

$f(x_i^1, x_j^2) \Delta x_i^1 \Delta x_j^2 + o(\Delta x_i^1 \Delta x_j^2)$ (дивись відповідну властивість двовимірної функції щільності).

Знайдемо математичне сподівання $M(\Psi(X_1^*, X_2^*))$. За формулою, введеною раніше $M(\Psi(X_1^*, X_2^*)) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \Psi(x_i^1, x_j^2) [f(x_i^1, x_j^2) \Delta x_i^1 \Delta x_j^2 + o(\Delta x_i^1 \Delta x_j^2)]$. Якщо існує інтеграл, модуль якого обмежений

$$\left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\Psi(x_1, x_2) f(x_1, x_2)) dx_1 dx_2 \right| < \infty,$$

то $\forall i, j \quad \Delta x_i^1, \Delta x_j^2 \rightarrow 0$, то існує границя для подвійної суми, що і є цим подвійним інтегралом.

Загальний випадок

Нехай $X_1, \dots, X_n$ – n -вимірна дискретна випадкова величина

$$M\Psi(X_1, \dots, X_n) = \sum_{j_1=1}^{S_1} \dots \sum_{j_n=1}^{S_n} \Psi(x_{j_1}^1, \dots, x_{j_n}^n) P(x_{j_1}^1, \dots, x_{j_n}^n)$$

Якщо $X_1, \dots, X_n$ – неперервні випадкові величини, то

$$M\Psi(X_1, \dots, X_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Psi(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n$$

Теорема. Математичне сподівання суми випадкових величин дорівнює сумі їх математичних сподівань.

$$\text{Тобто, якщо } \Psi(X_1, \dots, X_n) = \sum_{i=1}^n X_i, \text{ то } M \sum_{i=1}^n X_i = \sum_{i=1}^n M X_i.$$

ЛЕКЦІЯ 28

Нехай X_i – дискретні випадкові величини.

Доведено, що $M \sum_{i=1}^n X_i = \sum_{i=1}^n M X_i$

Розглянемо випадок двовимірної випадкової дискретної величини, ($n=2$).

Використовуємо формулу:

$$M\psi(X_1, X_2) = \sum_{j_1=1}^{S_1} \sum_{j_2=1}^{S_2} \psi(x_{j_1}^1, x_{j_2}^2) p(x_{j_1}^1, x_{j_2}^2), \text{ тоді}$$

$$M(X_1 + X_2) = \sum_{j_1=1}^{S_1} \sum_{j_2=1}^{S_2} (x_{j_1}^1 + x_{j_2}^2) p(x_{j_1}^1, x_{j_2}^2) = \sum_{j_1=1}^{S_1} x_{j_1}^1 \sum_{j_2=1}^{S_2} p(x_{j_1}^1, x_{j_2}^2) + \\ \sum_{j_2=1}^{S_2} x_{j_2}^2 \sum_{j_1=1}^{S_1} p(x_{j_1}^1, x_{j_2}^2) = M(X_1) + M(X_2).$$

Примітка: У другій подвійній сумі змінимо порядок сумування: спочатку по j_2 , потім по j_1 .

За 3-ю аксіомою Колмогорова:

$$\sum_{j_2=1}^{S_2} p(x_{j_1}^1, x_{j_2}^2) = P\left(\bigcup_{j_2}^{S_2} x_{j_1}^1 x_{j_2}^2\right) = P_{x_1}(x_{j_1}^1), \text{ оскільки } x_{j_2}^2 \text{ — довільне.}$$

$$\text{Аналогічно } \sum_{j_1=1}^{S_1} p(x_{j_1}^1, x_{j_2}^2) = P_{x_2}(x_{j_2}^2).$$

$$\text{Отже, } M(X_1 + X_2) = M(X_1) + M(X_2).$$

Розглянемо X_1, X_2 – двовимірну, неперервну, випадкову величину, ($n=2$).

Використовуємо формулу:

$$M\psi(X_1, X_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \psi(x_1, x_2) f(x_1, x_2) dx_1 dx_2. \text{ Тоді}$$

$$M(X_1 + X_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 + x_2) f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{+\infty} x_1 \left[\int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 \right] + \\ \int_{-\infty}^{+\infty} x_2 \left[\int_{-\infty}^{+\infty} f(x_1, x_2) dx_1 \right] = M(X_1) + M(X_2),$$

$$\text{де } \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 = f_{x_1}(x_1), \int_{-\infty}^{+\infty} f(x_1, x_2) dx_1 = f_{x_2}(x_2).$$

Примітка: У другому двократному інтегралі змінено порядок інтегрування.

$X_1 \dots X_n$ – n – вимірна, неперервна, випадкова величина, n – довільне:

$$M\psi(X_1 \dots X_n) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \psi(x_1 \dots x_n) f(x_1 \dots x_n) dx_1 \dots dx_n$$

$$\begin{aligned}
M \sum_{i=1}^n X_i &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (\sum_{i=1}^n x_i) f(x_1 \dots x_n) dx_1 \dots dx_n = \\
&\int_{-\infty}^{+\infty} x_1 [\int_{-\infty}^{+\infty} \underset{n-1}{\ddot{\cup}} \int_{-\infty}^{+\infty} f(x_1 \dots x_n) dx_2 \dots dx_n] dx_1 + \\
&\int_{-\infty}^{+\infty} x_2 [\int_{-\infty}^{+\infty} \underset{n-1}{\ddot{\cup}} \int_{-\infty}^{+\infty} f(x_1 \dots x_n) dx_1 dx_3 \dots dx_n] dx_2 + \dots + \\
&\int_{-\infty}^{+\infty} x_n [\int_{-\infty}^{+\infty} \underset{n-1}{\ddot{\cup}} \int_{-\infty}^{+\infty} f(x_1 \dots x_n) dx_1 \dots dx_{n-1}] dx_n = \sum_{i=1}^n M(X_i).
\end{aligned}$$

Примітка 1: У кожному інтегралі, окрім першого, відповідним чином змінюємо порядок інтегрування.

Примітка 2: Використаємо формулу минулої лекції для $k=1$

Примітка 3: Для дискретного випадку n -кратний інтеграл замінюється n -кратною сумою – довести самим.

Теорема: Математичне сподівання добутку випадкових величин дорівнює добутку їх математичних сподівань.

Доведення:

Для непервних випадкових величин.

$$\begin{aligned}
M(\prod_{i=1}^n X_i) &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \prod_{i=1}^n x_i \prod_{i=1}^n f_{x_i}(x_i) dx_1 \dots dx_n = \prod_{i=1}^n \int_{-\infty}^{+\infty} x_i f_{x_i}(x_i) dx_i = \\
&\prod_{i=1}^n M(X_i).
\end{aligned}$$

Коефіцієнт коваріації

Коефіцієнтом коваріації випадкових величин X, Y називається математичне сподівання:

$$\begin{aligned}
\text{cov}(X, Y) &= M[(X - MX)(Y - MY)] = M(XY - XMY - YMX + MXMY) = M(XY) - \\
&MXMY - MYMX + MXMY = M(XY) - MXMY.
\end{aligned}$$

Отримано на основі наступної формули:

$$\text{Нехай } \psi(X_1 \dots X_n) = \sum_{l=1}^k \psi_l(X_{l_1} \dots X_{l_{k_l}}),$$

$$\text{тоді } M(\sum_{l=1}^k \psi_l(X_{l_1} \dots X_{l_{k_l}})) = \sum_{l=1}^k M\psi_l(X_{l_1} \dots X_{l_{k_l}}).$$

Примітка. Для нашого випадку $k=4, n=2$.

$$\psi(X, Y) = XY - XMY - YMX + MXMY,$$

$$\psi_1 = XY, \psi_2 = -XMY, \psi_3 = -YMX, \psi_4 = MXMY.$$

Доведемо цю формулу для неперервного випадку, для дискретного – аналогічно.

$$M\left(\sum_{l=1}^n \psi_l(x_{l_1} \dots x_{l_{k_l}})\right) = \int_{-\infty}^{+\infty} \ddot{\int}_{n-k_l}^{+\infty} \sum \psi_l(x_{l_1} \dots x_{l_{k_l}}) f(x_1 \dots x_n) dx_1 \dots dx_n$$

$$= \sum_{l=1}^k \int_{-\infty}^{+\infty} \ddot{\int}_{k_l}^{+\infty} \psi_l(x_{l_1} \dots x_{l_{k_l}}) \left[\int_{-\infty}^{+\infty} \ddot{\int}_{n-k_l}^{+\infty} f(x_1 \dots x_n) dx_{j_1} \dots dx_{j_{n-k_l}} \right] dx_{l_{k_l}} \dots dx_{l_1}.$$

Примітка: В кожному n-кратному інтегралі, що входить в суму, відповідним чином змінюємо порядок інтегрування: спочатку йдуть зовнішні інтеграли по змінним $dx_{l_{k_l}} \dots dx_{l_1}$, а у внутрішніх - всі інші.

Використовується формула:

$$\left[\int_{-\infty}^{+\infty} \ddot{\int}_{n-k_l}^{+\infty} f(x_1 \dots x_n) dx_{j_1} \dots dx_{j_{n-k_l}} \right] = f_{l_1 \dots l_{k_l}}(x_{l_1} \dots x_{l_{k_l}}),$$

таким чином:

$$M\left(\sum_{l=1}^n \psi_l(X_{l_1} \dots X_{l_{k_l}})\right)$$

$$= \sum_{l=1}^k \int_{-\infty}^{+\infty} \ddot{\int}_{k_l}^{+\infty} \psi_l(x_{l_1} \dots x_{l_{k_l}}) f_{l_1 \dots l_{k_l}}(x_{l_1} \dots x_{l_{k_l}}) dx_{l_{k_l}} \dots dx_{l_1}$$

$$= \sum_{l=1}^n M\psi_l(X_{l_1} \dots X_{l_{k_l}}).$$

Таким чином

$$\text{cov}(X, Y) = M[(X - MX)(Y - MY)] = M(XY) - MXMY$$

Наслідок 1: Якщо X і Y незалежні, то $M(XY) = MXMY \Rightarrow \text{cov}(X, Y) = 0$

Наслідок 2: З того, що коваріація випадкових величин дорівнює нулю, не впливає, що вони незалежні:

Це впливає з наступного прикладу:

Розглянемо випадкову величину X, що має нормальний розподіл з нульовим математичним сподіванням. Нехай $Y=X^2$. Очевидно, що X та Y – залежні випадкові величини та $MX=0$.

Покажемо, що $\text{cov}(X, Y) = 0$

$$\text{cov}(X, X^2) = M[X(X^2 - MX^2)] = MX^3 - MXMX^2 = 0,$$

оскільки раніше було доведено, що якщо X – розподілена нормально і $MX=0$, то всі непарні початкові моменти дорівнюють нулю.

Таким чином, маємо дві випадкові величини, які мають нульову коваріацію, але є залежними.

ЛЕКЦІЯ 29

Коефіцієнт кореляції

Нехай X – довільна випадкова величина.

Означення. Нормованою випадковою величиною X називають випадкову величину X^* :

$$X^* = \frac{X - v_X}{\sigma_X},$$

де $v_X = MX$, $\sigma_X = \sqrt{DX}$.

Властивості нормованої випадкової величини X^* :

1) $MX^* = 0$,

2) $DX^* = 1$.

Примітка. Довести самостійно, використовуючи властивості математичного сподівання та дисперсії.

Означення. Коефіцієнтом кореляції випадкових величин X та Y називають коваріацію величин X^* та Y^* , тобто

$$\rho_{XY} = \text{cov}(X^*, Y^*) = M \left[\frac{X - v_X}{\sigma_X} \cdot \frac{Y - v_Y}{\sigma_Y} \right] = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \Rightarrow$$
$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Таким чином, коефіцієнт кореляції є часткою від ділення коваріації на добуток середніх квадратичних відхилень випадкових величин.

Властивості коефіцієнта кореляції ρ_{XY}

Розглянемо наступні формули

$$D(X \pm Y) = M[(X \pm Y) - M(X \pm Y)]^2 = M[(X - MX) \pm (Y - MY)]^2 = \\ = M(X - MX)^2 + M(Y - MY)^2 \pm 2M[(X - MX)(Y - MY)] = DX + DY \pm 2\text{cov}(X, Y).$$

$$D(X \pm Y) = DX + DY \pm 2\text{cov}(X, Y).$$

Наслідок. Якщо X та Y – незалежні, то $\text{cov}(X, Y) = 0$ і

$$D(X \pm Y) = DX + DY.$$

$$D\left(\sum_{i=1}^n \pm X_i\right) = \sum_{i=1}^n D(X_i) + \sum_{i=1}^n \sum_{j=1}^n \pm \text{cov}(X_i, X_j).$$

Дисперсія суми (різниці) незалежних випадкових величин дорівнює сумі (різниці) дисперсій.

Розглянемо дисперсію: $D(X^* \pm Y^*) = 2 \pm 2\rho_{XY} \geq 0$. Звідси маємо

1. $-1 \leq \rho_{XY} \leq 1$
2. $\rho_{XY} = 0$, якщо випадкові величини X та Y є незалежними.
3. Якщо $|\rho_{XY}| = 1$, то випадкові величини X та Y мають лінійну залежність.

Доведення.

а) $\rho_{XY} = -1$

$$M(X^* + Y^*) = 0, \quad D(X^* + Y^*) = 0$$

$X^* + Y^* = 0$ (з ймовірністю 1), отже, матимемо

$$X^* = -Y^*$$

$$\frac{X - v_X}{\sigma_X} = -\frac{Y - v_Y}{\sigma_Y},$$

$$Y = v_Y - \frac{\sigma_Y}{\sigma_X}(X - v_X),$$

$$Y = a_1 X + b_1,$$

$$a_1 = -\frac{\sigma_Y}{\sigma_X},$$

$$b_1 = v_Y + \frac{\sigma_Y}{\sigma_X} v_X,$$

б) $\rho_{XY} = 1$, тоді

$$M(X^* - Y^*) = 0, \quad D(X^* - Y^*) = 0$$

$$\frac{X - v_X}{\sigma_X} = \frac{Y - v_Y}{\sigma_Y},$$

$$Y = v_Y + \frac{\sigma_Y}{\sigma_X} (X - v_X),$$

$$Y = a_2 X + b_2$$

$$a_2 = \frac{\sigma_Y}{\sigma_X},$$

$$b_2 = v_Y - \frac{\sigma_Y}{\sigma_X} v_X,$$

4. Коефіцієнт кореляції є якісною мірою лінійного зв'язку між довільними випадковими величинами X та Y , а саме:

чим ближче по модулю коефіцієнт кореляції ρ_{XY} до одиниці – тим тісніше лінійний зв'язок між випадковими величинами X та Y .

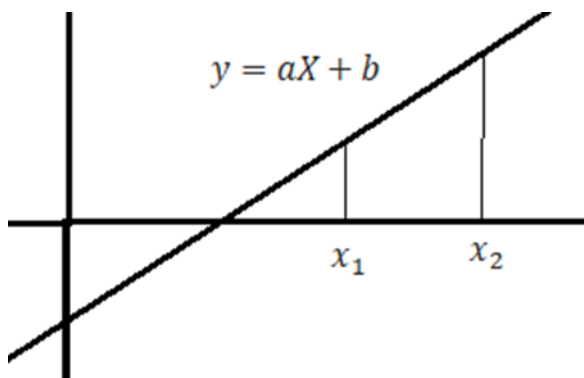
Нехай $\rho_{XY} \neq 0$. Доводиться, що має місце наступний результат

$$Y = aX + b + Z, \quad \text{де } Z \text{ – випадкова величина}$$

$$MZ = 0, \quad DZ = \sigma_Y^2(1 - \rho_{XY}^2)$$

Примітка. Нехай $\rho_{XY} < 0$, тоді $a = a_1$, $b = b_1$.

Нехай $\rho_{XY} > 0$, тоді $a = a_2$, $b = b_2$.



$$Y = aX + b + Z$$

Дійсно, якщо $|\rho_{XY}|$ близький до одиниці, то дисперсія випадкової величини Z мала, а її математичне сподівання дорівнює ($MZ = 0$), тому результати випробування над Z малі за модулем. Тоді, якщо на площині відкласти

результати випробувань над випадковими величинами X та Y , отримаємо множину точок, що тісно групуються відносно прямої $y = ax + b$.

Постановка задачі: X, Y – незалежні неперервні випадкові величини. З означення випливає, що їх двовимірна функція щільності має вигляд

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

Знайти функцію щільності випадкової величини $Z = X + Y$.

Спочатку знайдемо вираз для функції розподілу випадкової величини Z , а потім візьмемо похідну по z .

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) \quad (P((X, Y) \in G) = \iint_G f(x, y) dx dy).$$

Подія $X + Y \leq z$ складається з будь-яких пар чисел x та y , таких що $x + y \leq z$.

Тому використовуємо загальну формулу для випадку $G = x + y \leq z$, отримаємо:

$$\begin{aligned} F_Z(z) &= P(x + y \leq z) = \iint_{x+y \leq z} f(x, y) dx dy = \iint_{x+y \leq z} f_X(x) \cdot f_Y(y) dx dy = \\ &= \int_{-\infty}^{+\infty} f_X(x) \left[\int_{-\infty}^{z-x} f_Y(y) dy \right] dx = \int_{-\infty}^{+\infty} f_Y(y) \left[\int_{-\infty}^{z-y} f_X(x) dx \right] dy \\ f_Z(z) &= \frac{dF_Z(z)}{dz} = \int_{-\infty}^{+\infty} f_X(x) \left[\frac{d}{dz} \int_{-\infty}^{z-x} f_Y(y) dy \right] dx = \int_{-\infty}^{+\infty} f_Y(y) \left[\frac{d}{dz} \int_{-\infty}^{z-y} f_X(x) dx \right] dy \end{aligned}$$

Примітка.

$$\frac{d}{dx} \int_{-\infty}^x f(u) du = f(x).$$

Покажемо, що

$$\frac{d}{dx} \int_{-\infty}^{x-a} f(u) du = f(x - a)$$

Робимо заміну $v = u + a$, тоді

$$\frac{d}{dx} \int_{-\infty}^{x-a} f(u) du = \frac{d}{dx} \int_{-\infty}^x f(v - a) dv = f(x - a).$$

Отже,

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty}^{+\infty} f_Y(y)f_X(z-y)dy.$$

Двовимірний нормальний розподіл

Означення. Двовимірна неперервна випадкова величина (X, Y) має двовимірний нормальний розподіл, якщо її функція щільності має вигляд:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp \left\{ -\frac{1}{2(1-\rho_{XY}^2)} \left[\frac{(x-v_X)^2}{\sigma_X^2} + \frac{(y-v_Y)^2}{\sigma_Y^2} - 2\rho_{XY} \cdot \frac{(x-v_X)}{\sigma_X} \cdot \frac{(y-v_Y)}{\sigma_Y} \right] \right\},$$

де $v_X = MX$, $v_Y = MY$, $\sigma_X = +\sqrt{DX}$, $\sigma_Y = +\sqrt{DY}$.

ЛЕКЦІЯ 30

Властивості двовимірного нормального розподілу

1. $f(x, y)$ задовольняє всім характеристичним властивостям функції щільності: неперервна, невід'ємна та справедлива рівність

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$

$$2. \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-v_X)^2}{2\sigma_X^2}},$$

$$\int_{-\infty}^{+\infty} f(x, y) dx = \frac{1}{\sigma_Y \sqrt{2\pi}} e^{-\frac{(y-v_Y)^2}{2\sigma_Y^2}},$$

оскільки X і Y мають нормальний розподіл.

3. Якщо $\rho_{XY} = 0$, то X та Y незалежні і двовимірна функція щільності дорівнює добутку двох одновимірних.

Зворотнє твердження невірне: якщо X і Y розподілено нормально, $\rho_{XY} = 0$, тоді X та Y – незалежні – НЕ ВІРНО.

Знайдемо умовну функцію щільності:

$$f_y(x/y) = \frac{f(x,y)}{f_x(x)}$$

$$u = \frac{x-v_x}{\sigma_x}, v = \frac{y-v_y}{\sigma_y}, \text{ тоді}$$

$$f_y(x/y) = f(x,y) = \frac{\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} \left[e^{\frac{-1}{2(1-\rho_{xy}^2)}(u^2+v^2-2\rho_{xy}uv)} \right]}{\frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{u^2}{2}}} =$$

Примітка: замість x, y підставляємо u та v в чисельник і знаменник

$$= \frac{1}{\sigma_y\sqrt{2\pi}\sqrt{1-\rho_{xy}^2}} \left[e^{\frac{-1}{2(1-\rho_{xy}^2)}(v^2-2v\rho_{xy}u+\rho_{xy}^2u^2)} \right] = \frac{1}{\sigma_y\sqrt{2\pi}\sqrt{1-\rho_{xy}^2}} \left[e^{\frac{-1}{2(1-\rho_{xy}^2)}(v-\rho_{xy}u)^2} \right] =$$

$$\frac{1}{\sigma_y\sqrt{2\pi}\sqrt{1-\rho_{xy}^2}} \left[e^{\frac{-\sigma_y^2}{2\sigma_y^2(1-\rho_{xy}^2)}\left(\frac{y-v_x}{\sigma_y}-\rho_{xy}\frac{x-v_x}{\sigma_x}\right)^2} \right] =$$

$$\frac{1}{\sigma_y\sqrt{2\pi}\sqrt{1-\rho_{xy}^2}} \left[e^{\frac{-1}{2\sigma_y^2(1-\rho_{xy}^2)}\left(y-v_x-\rho_{xy}\frac{\sigma_y}{\sigma_x}(x-v_x)\right)^2} \right] = f(x/y).$$

$$\bar{y}(x) = v_x + \rho_{xy} \frac{\sigma_y}{\sigma_x} (x - v_x).$$

$$\sigma_{y/x}^2 = \sigma_y^2 (1 - \rho_{xy}^2).$$

Висновки

1. Умовна функція щільності має вигляд нормального розподілу, в якому замість безумовного математичного сподівання стоїть умовне, а замість безумовної дисперсії – умовна.
2. Умовне математичне сподівання – лінійна функція від x .
3. Умовна дисперсія від x не залежить і є константою.

***Багатовимірний нормальний розподіл

Неперервна багатовимірна випадкова величина $X_1 \dots X_n$ має n -вимірний нормальний розподіл, якщо

$$f(x_1 \dots x_n) = n(x, v_x, B) = \frac{|B|^{\frac{-1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(x-v_x)^T B^{-1}(x-v_x)}.$$

$x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ – векторний аргумент.

$v_x = \begin{pmatrix} MX_1 \\ \dots \\ MX_n \end{pmatrix}$ – вектор, координати якого математичні сподівання випадкової величини.

$B = (ij)_1^n$ – коваріаційна матриця n – вимірної випадкової величини $X_1 \dots X_n$

$(ij) = \text{cov}(X_i, X_j)$

Примітка: На головній діагоналі коваріаційної матриці стоять дисперсії випадкових величин X_i .

$|B|$ – детермінант B

B^{-1} – обернена матриця

Перевіряємо характеристичні властивості функції $n(x, v_x, B)$.

1) Це числова скалярна функція n дійсних аргументів.

Властивості:

1. Неперервна
2. Невід’ємна.

Щоб довести невід’ємність, треба довести, що $|B| \geq 0$.

Довільна квадратна матриця B називається **невід’ємно-означеною**, якщо для будь-якого дійсного вектора u квадратична форма $u^T B u$ є невід’ємним числом.

$$u^T B u \geq 0 \Rightarrow \sum_{i=1}^n \sum_{j=1}^n u_i B_{ij} u_j.$$

Наслідок: Детермінанти всіх головних мінорів B – невід’ємні числа.

ЛЕКЦІЯ 31

Покажемо, що матриця B є невід’ємно-означеною, тобто що $u^T B u \geq 0$.

$$B = \begin{pmatrix} M[(X_1 - MX_1)(X_1 - MX_1)] & \dots & M[(X_1 - MX_1)(X_n - MX_n)] \\ \vdots & \dots & \vdots \\ M[(X_i - MX_i)(X_1 - MX_1)] & \ddots & M[(X_i - MX_i)(X_n - MX_n)] \\ \vdots & \dots & \vdots \\ M[(X_n - MX_n)(X_1 - MX_1)] & \dots & M[(X_n - MX_n)(X_n - MX_n)] \end{pmatrix}$$

Беремо довільний вектор $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Спочатку перемножимо вектор-рядок

$(y_1 \dots y_n)$ на матрицю B , а потім отриманий результат помножимо на вектор-стовпець $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Позначимо $Y_i = y_i X_i$. Матимемо,

$$\sum_{i=1}^n M[(Y_i - MY_i)(X_1 - MX_1)] \dots \sum_{i=1}^n M[(Y_i - MY_i)(X_n - MX_n)].$$

Тепер множимо отриманий вектор-рядок на вектор стовпець $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, отримаємо:

$$y^T B y = \sum_{i=1}^n \sum_{j=1}^n M[(Y_i - MY_i)(Y_j - MY_j)].$$

Примітка: числа y_i вносяться під другу дужку кожного доданка і робиться знову заміна $Y_i = y_i X_i$. Використаємо властивість математичного сподівання: математичне сподівання суми дорівнює сумі математичних сподівань.

$$\begin{aligned} y^T B y &= M \sum_{i=1}^n \sum_{j=1}^n [(Y_i - MY_i)(Y_j - MY_j)] = \left[\sum_{i=1}^n \sum_{j=1}^n a_i a_j = \left(\sum_{i=1}^n a_i \right)^2 \right] = \\ &= \left[\sum_{i=1}^n M(Y_i - MY_i) \right]^2 \geq 0. \end{aligned}$$

Таким чином, матриця B є невід'ємно-означеною.

Властивості (аналог властивості двовимірного випадку без доведення):

3) При двовірному випадку

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} n(x, v_X, B) dx_1 \dots dx_n = 1$$

4) Нехай $k < n$, $X_{i_1}, \dots, X_{i_k}$ – k -вимірний випадковий розмір, тоді

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} n(x, v_X, B) dx_{j_1} \dots dx_{j_{n-k}} = f(x_{i_1} \dots x_{i_k}) = n(x, v_X, B).$$

Примітка: Отримали k -вимірну функцію щільності (доведено раніше для будь-якої функції щільності).

$$x = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_k} \end{pmatrix}, \quad v_X = \begin{pmatrix} MX_{i_1} \\ \vdots \\ MX_{i_k} \end{pmatrix}, \quad B = (\text{cov}(X_{i_l}, X_{i_m}))^k$$

Тобто, ця k -вимірна випадкова величина має k -вимірний нормальний розподіл.

Наслідок. Кожна випадкова величина X_i має нормальний розподіл

$$X_i = n(x_i, v_{X_i}, \sigma_{X_i}) = \frac{1}{\sqrt{2\pi}\sigma_{X_i}} e^{-\frac{(x_i - v_{X_i})^2}{2\sigma_{X_i}^2}}.$$

5) щільність розподілу має вигляд

$$\frac{|B|^{-\frac{1}{2}}}{(2\pi)^{n/2}} e^{-\frac{1}{2}(x-v_X)^T B^{-1}(x-v_X)} = n(x, v_X, B).$$

Нехай $\forall \rho_{x_i x_j} = 0, i \neq j$, тоді B – діагональна матриця і B^{-1} – теж діагональна. В цьому випадку багатовимірна функція щільності дорівнює добутку одновимірних нормальних функцій щільності, тобто X_i – незалежні.

Обернене твердження невірне: тобто з того, що X_i мають нормальний розподіл, $\forall \rho_{x_i x_j} = 0, i \neq j$, не випливає, що X_i – незалежні (оскільки $X_1 \dots X_n$ не обов'язково мають n -вимірний нормальний розподіл).

***Теорема.

Нехай вектор $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ має n -вимірний нормальний розподіл. Матриця A –

невироджена квадратна матриця розмірності $n \times n$ ($|A| \neq 0$). $Y = AX$. Тоді

компоненти вектора $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ мають n -вимірний нормальний розподіл

$n(x, Av_X, ABA^T)$, тобто вектор математичних сподівань $\begin{pmatrix} MY_1 \\ \vdots \\ MY_n \end{pmatrix} = Av_X$ і коваріаційна

матриця компонент вектора Y дорівнює ABA^T .

Доведення.

Покажемо, що $\begin{pmatrix} MY_1 \\ \vdots \\ MY_n \end{pmatrix} = Av_X$.

Нехай C – довільна матриця з випадковими елементами, тоді $MC = (M c_{ij})_1^n$.

$$Y=AX \Rightarrow MY=MAX=AMX=Av_X.$$

Покажемо, що коваріаційна матриця компонент вектора Y дорівнює ABA^T .

Нехай B – коваріаційна матриця випадкових величин $X_1 \dots X_n$, тоді

$$B = \begin{bmatrix} (X_1 - MX_1) \\ \vdots \\ (X_n - MX_n) \end{bmatrix} ((X_1 - MX_1) \dots (X_n - MX_n)) = M[(X - MX)(X - MX)^T].$$

Запишемо коваріаційну матрицю компонент Y :

$$\begin{aligned} M[(Y - MY)(Y - MY)^T] &= M[(AX - AMX)(AX - AMX)^T] = \\ &= |\text{використаємо формулу } (AB)^T = B^T A^T| = \\ &= M[A(X - MX)(X - MX)^T A^T] = A \cdot M[(X - MX)(X - MX)^T] A^T = ABA^T. \end{aligned}$$

Наслідок.

Такий само результат буде для довільної прямокутної матриці A (в доведенні ніде не використовується факт, що A – квадратна) і для довільного розподілу $X_1 \dots X_n$ (в доведенні ніде не вказувалось, що $X_1 \dots X_n$ нормальний розподіл).

Покажемо, що $Y_1 \dots Y_n$ – має n -вимірний нормальний розподіл, тобто функція щільності має вигляд

$$f(y_1 \dots y_n) = n(y, Av_X, ABA^T) = \frac{|ABA^T|^{-1/2}}{(2\pi)^{n/2}} e^{-\frac{1}{2}(y-Av_X)^T(ABA^T)^{-1}(y-Av_X)}.$$

ЛЕКЦІЯ 32

Розглянемо довільну область $S \in R^n$.

$$P\left(\underbrace{X_1 \dots X_n}_{\text{вектор рядок}} \in S\right) = \iint_S n(x, v_x, B) dx_1 \dots dx_n =$$

Знайдемо $P(Y_1 \dots Y_n \in A(S))$.

Оскільки Y і X зв'язані функціонально ($Y=AX$), то якщо X приймає значення з області S , то Y приймає значення з області $A(S)$.

Примітка: область $A(S)$ складається з усіх векторів x , $\forall y = Ax, \forall x \in S$. Це означає ймовірність, що $X \in S$ така сама як і ймовірність $Y \in A(S)$, $P(Y_1 \dots Y_n \in A(S))$

$$= \iint_{A(S)} f(y_1 \dots y_n) dy_1 \dots dy_n.$$

В першому інтегралі зробимо заміну змінних:

$$x = A^{-1}y = \begin{pmatrix} a_{11}^{-1} & \dots & a_{1i}^{-1} & \dots & a_{1n}^{-1} \\ \dots & & \dots & & \dots \\ a_{i1}^{-1} & \dots & a_{ii}^{-1} & \dots & a_{in}^{-1} \\ \dots & & \dots & & \dots \\ a_{n1}^{-1} & \dots & a_{ni}^{-1} & \dots & a_{nn}^{-1} \end{pmatrix} y =$$

$$= \begin{pmatrix} a_{11}^{-1}y_1 + \dots + a_{1i}^{-1}y_i + \dots + a_{1n}^{-1}y_n \\ \dots \\ a_{i1}^{-1}y_1 + \dots + a_{ii}^{-1}y_i + \dots + a_{in}^{-1}y_n \\ \dots \\ a_{n1}^{-1}y_1 + \dots + a_{ni}^{-1}y_i + \dots + a_{nn}^{-1}y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \dots \\ x_i \\ \dots \\ x_n \end{pmatrix}.$$

Тоді перший інтеграл матиме вигляд:

$$\iint_S n(x, v_x, B) dx_1 \dots dx_n = \iint_{A(S)} n(A^{-1}y, v_x, B) \det(J) dy_1 \dots dy_n.$$

$$\text{де } J = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \dots & \frac{\partial x_1}{\partial y_i} & \dots & \frac{\partial x_1}{\partial y_n} \\ \dots & & \dots & & \dots \\ \frac{\partial x_i}{\partial y_1} & \dots & \frac{\partial x_i}{\partial y_i} & \dots & \frac{\partial x_i}{\partial y_n} \\ \dots & & \dots & & \dots \\ \frac{\partial x_n}{\partial y_1} & \dots & \frac{\partial x_n}{\partial y_i} & \dots & \frac{\partial x_n}{\partial y_n} \end{pmatrix} = \begin{pmatrix} a_{11}^{-1} & \dots & a_{1i}^{-1} & \dots & a_{1n}^{-1} \\ \dots & & \dots & & \dots \\ a_{i1}^{-1} & \dots & a_{ii}^{-1} & \dots & a_{in}^{-1} \\ \dots & & \dots & & \dots \\ a_{n1}^{-1} & \dots & a_{ni}^{-1} & \dots & a_{nn}^{-1} \end{pmatrix},$$

$$\det(J) = \det(A^{-1}).$$

Оскільки S – довільна область, отримано результат:

$$f(y_1 \dots y_n) = n(A^{-1}y, v_x, B) |A^{-1}| = |A|^{-1} \frac{|B|^{-1/2}}{(2\pi)^{n/2}} e^{-\frac{1}{2}(A^{-1}y - v_x)^T B^{-1} (A^{-1}y - v_x)}.$$

Рівність виразів під інтегралами впливає з того, що область S – довільна. Далі використовується наступний список формул:

$$(AB)^T = B^T A^T$$

$$|A| = |A^T|$$

$$|AB| = |A||B|$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A^{-1})^T = (A^T)^{-1}$$

$$|A^{-1}A| = |A^{-1}||A| = 1 \Rightarrow |A^{-1}| = \frac{1}{|A|} = \sqrt{\frac{|B|}{|A||B||A^T|}} = |B|^{1/2}|ABA^T|^{-1/2}.$$

$$\text{Тому } |A|^{-1} \frac{|B|^{-1/2}}{(2\pi)^{n/2}} = \frac{|ABA^T|^{-1/2}}{(2\pi)^{n/2}}.$$

Перетворимо вираз $e^{(A^{-1}y - v_x)^T B^{-1}(A^{-1}y - v_x)}$.

Покажемо, що він дорівнює: $e^{(y - Av_x)^T (ABA^T)^{-1}(y - Av_x)}$.

$$\begin{aligned} (A^{-1}y - v_x)^T B^{-1}(A^{-1}y - v_x) &= (A^{-1}y - A^{-1}Av_x)^T B^{-1}(A^{-1}y - A^{-1}Av_x) = \\ &= (A^{-1}(y - Av_x))^T B^{-1}A^{-1}(y - Av_x) = (y - Av_x)^T (A^T)^{-1}(AB)^{-1}(y - Av_x) \\ &= (y - Av_x)^T (ABA^T)^{-1}(y - Av_x). \end{aligned}$$

Наслідок: Нехай матриця A – прямокутна $m \times n, m < n$ і ранг матриці $A=m$ ($\text{rang}(A)=m$), тобто матриця A має m лінійно-незалежних стовпців.

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_m \end{pmatrix} = AX = A \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

Нехай компоненти вектора X мають n -вимірний нормальний розподіл, тоді $(Y_1 \dots Y_m)$ -мають m -вимірний нормальний розподіл.

Доведення.

Випадки

а) Розглянемо спочатку частковий випадок, коли матриця A має блочну структуру $A=(A_1 A_2)$, де матриця A_1 вимірності $m \times m$ не вироджена.

Виведемо матрицю

$$C = \begin{pmatrix} A_1 & A_2 \\ 0 & E_{n-m} \end{pmatrix},$$

де E_{n-m} – одинична матриця вимірності $(n-m) \times (n-m)$. Матриця C не вироджена, тому що

$$|C| = |A_1| |E_{n-m}| = |A_1| \neq 0.$$

Як було показано вище компоненти випадкового вектора Z ,

$$Z = CX = \begin{pmatrix} A_1 & A_2 \\ 0 & E_{n-m} \end{pmatrix} \begin{pmatrix} X_1 \\ \dots \\ X_n \end{pmatrix} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_m \\ X_{m+1} \\ \vdots \\ X_n \end{pmatrix}$$

мають n – вимірний нормальний розподіл, при чому перші m його компонентів складають вектор Y . Але ми знаємо, що коли вектор має n – вимірний нормальний розподіл, то і довільна сукупність його компонент має відповідний багатовимірний нормальний розподіл. Тобто в цій ситуації твердження має місце .

б) Загальний випадок

m -лінійно незалежних стовпців матриці A стоять довільно.

$$Y = AX$$

Переставимо в матриці A m -лінійно незалежних стовпців на перші місця і, відповідним чином, пронумеруємо компоненти вектора X :

$$AX = (A_1 A_2) X_1,$$

де квадратна матриця A_1 – невироджена, а компоненти вектора $X_1 = \begin{pmatrix} X_{i1} \\ \vdots \\ X_{in} \end{pmatrix}$

відповідають новій нумерації стовпців матриці $(A_1 A_2)$.

$$Y = AX = (A_1 A_2) X_1.$$

Примітка:

- компоненти вектора X_1 мають n -вимірний нормальний розподіл,
- в функції щільності вектора X потрібно поміняти відповідним чином порядок аргументів.

Отримали випадок а).

Приклад: $X_1 \dots X_n$ мають n-вимірний нормальний розподіл, за яким законом розподілена випадкова величина:

$$Y = \sum_{i=1}^n a_i X_i = (a_1 \dots a_n) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

Відповідь: Y має одновимірний нормальний розподіл.

Перший параметр цього розподілу – математичне сподівання $\sum_{i=1}^n a_i X_i$, другий – дисперсія $\sum_{i=1}^n a_i^2 X_i$.

ЛЕКЦІЯ 33

***Збіжність випадкових послідовностей

Маємо ймовірнісний простір (Ω, σ, P) , на якому задається послідовність випадкових величин

$$\varphi_1(\omega), \varphi_2(\omega), \dots, \varphi_n(\omega), \dots$$

Задамо подію

$$A = \left\{ \omega : \bigcap_{r=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} \left(|\varphi_n(\omega) - \varphi_m(\omega)| \leq \frac{1}{r} \right) \right\}.$$

Означення 1. Випадкова послідовність $\varphi_1(\omega), \varphi_2(\omega), \dots, \varphi_n(\omega), \dots$ збігається з ймовірністю 1 до випадкової величини $\varphi(\omega)$, якщо $P(A) = 1$.

Примітка 1. $P(A)$ існує, тому що A є елементом δ -алгебри і тому має ймовірність настання.

Примітка 2. Подія A складається з усіх елементарних подій ω , для кожної з яких виконується: для будь-якого як завгодно великого, але обмеженого натурального r існує таке натуральне m , що для всіх $n \geq m$ виконується нерівність:

$$|\varphi_n(\omega) - \varphi_m(\omega)| \leq \frac{1}{r}.$$

Еквівалентне означення.

$$A = \left\{ \omega : \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} \left(|\varphi_n(\omega) - \varphi(\omega)| \leq \frac{1}{r} \right) \right\}, \text{ де } \varphi(\omega) - \text{гранична випадкова величина.}$$

Означення 2 (збіжність за ймовірністю). Послідовність випадкових величин $\varphi_n(\omega)$ збігається за ймовірністю до випадкової величини або константи, якщо для кожного $r > 0$

$$\lim_{n \rightarrow \infty} P \left\{ \omega : \left| \varphi_n(\omega) - (\varphi(\omega) \vee h) \right| > \frac{1}{r} \right\} = 0.$$

Примітка. Далі буде показано, що збіжність за ймовірністю більш «повільна», ніж збіжність з ймовірністю 1.

Означення 3. Послідовність випадкових величин збігається в середньоквадратичному до випадкової величини $\varphi(\omega) \vee h$, якщо існує наступна границя:

$$\lim_{n \rightarrow \infty} M \left(\varphi_n(\omega) - (\varphi(\omega) \vee h) \right)^2 = 0, \text{ чи}$$

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} M \left(\varphi_n(\omega) - \varphi_m(\omega) \right)^2 = 0.$$

Наслідок. З збіжності в середньоквадратичному випливає збіжність за ймовірністю (збіжність в середньоквадратичному не повільніша за збіжність за ймовірністю).

Доведення. Для доведення цього наслідку використаємо нерівність Чебишова:

$$P(Z \geq \tau) \leq \frac{MZ}{\tau}, \quad Z \geq 0, \tau > 0;$$

В даному випадку $P(|\varphi_n(\omega) - \varphi(\omega)| \geq \frac{1}{r}) = P(|\varphi_n(\omega) - \varphi(\omega)|^2 \geq \frac{1}{r})$

отже

$$0 \leq P \left(\omega : \left| \varphi_n(\omega) - \varphi(\omega) \right| > \frac{1}{r} \right) \leq \frac{M \left[\varphi_n(\omega) - \varphi(\omega) \right]^2}{\left(\frac{1}{r} \right)^2}.$$

Знаменник у правій частині скільки завгодно малий, але більший нуля. Тому, коли чисельник прямує до нуля, то і вся права частина прямує до нуля, а, тим паче, ліва частина нерівності.

Отже, довели що із збіжності в середньоквадратичному випливає збіжність за ймовірністю.

Нехай на ймовірнісному просторі (Ω, σ, P) задано послідовність випадкових величин

$$\varphi_1(\omega), \varphi_2(\omega), \dots, \varphi_n(\omega).$$

Ці n випадкових величин породжують новий простір $(R_n, \mathcal{B}_n, P_{F(x_1 \dots x_n)})$ (в цьому випадку випадкові величини позначаються як $X_1, X_2, \dots, X_n$).

Маємо еквівалентне означення всіх трьох введених збіжностей:

1) збіжність з ймовірністю 1

$$P\left(\bigcap_{r=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} \left(|X_n - X_m| \leq \frac{1}{r}\right)\right) = 1,$$

$$P\left(|X_n - (X \vee h)| \leq \frac{1}{r}\right) = 1.$$

2) збіжність за ймовірністю

$$\exists \lim_{n \rightarrow \infty} P\left(|X_n - (X \vee h)| > \frac{1}{r}\right) = 0,$$

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} P\left(|X_n - X_m| > \frac{1}{r}\right) = 0.$$

$$\exists \lim_{n \rightarrow \infty} P(|X_n - (X \vee h)| > \frac{1}{r}) = 0$$

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} P(|X_n - X_m| > \frac{1}{r}) = 0$$

3) збіжність в середньоквадратичному

$$\exists \lim_{n \rightarrow \infty} M(X_n - (X \vee h))^2 = 0,$$

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} M(X_n - X_m)^2 = 0.$$

Теорема Бернуллі

Розглянемо схему Бернуллі: проводиться n незалежних випробувань, у кожному з яких подія A відбувається з імовірністю p , та не відбувається (відбувається подія $\bar{A}$) з імовірністю $q=1-p$. З кожним i -им випробуванням пов'яжемо дискретну випадкову величину X_i , $i = \overline{1, n}$, таку що

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

X_i приймає значення 0, якщо в i -ому випробуванні настала подія $\bar{A}$ та 1, якщо настала подія A .

Випадкові величини X_i – незалежні, оскільки незалежними є випробування Бернуллі.

Математичне сподівання, другий початковий момент, дисперсія цієї дискретної випадкової величини дорівнюють:

$$MX_i = 1 \cdot p + 0 \cdot (1-p) = p;$$

$$MX_i^2 = 1 \cdot p + 0 \cdot (1-p) = p;$$

$$DX_i = MX_i^2 - (MX_i)^2 = p - p^2 = p(1-p) = pq.$$

Розглянемо частоту настання випадкової події A в n незалежних випробуваннях, тобто випадкову величину $S_n = \sum_{i=1}^n X_i$. За властивостями математичного сподівання та дисперсії суми незалежних випадкових величин $X_i, i = \overline{1, n}$, маємо:

$$MS_n = M \sum_{i=1}^n X_i = \sum_{i=1}^n MX_i = np;$$

$$DS_n = D \sum_{i=1}^n X_i = \sum_{i=1}^n DX_i = npq.$$

Для частоти настання події A в n незалежних випробуваннях, тобто для випадкової величини $\frac{S_n}{n}$, маємо:

$$M\left(\frac{S_n}{n}\right) = p, \quad D\left(\frac{S_n}{n}\right) = \frac{1}{n^2} npq = \frac{pq}{n}.$$

За нерівністю Чебишова для довільного $\varepsilon > 0$

$$P\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) \leq \frac{pq}{n\varepsilon^2},$$

Примітка. Використовується нерівність Чебишова у вигляді

$$P(|X - MX| \geq \varepsilon) \leq \frac{DX}{\varepsilon^2}, \forall \varepsilon > 0.$$

і, таким чином,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) = 0.$$

Ми довели наступний дуже важливий результат.

Теорема Бернуллі. Частість настання події в n незалежних випробуваннях збігається за ймовірністю до аксиоматичної за Колмогоровим ймовірності цієї події.

Нехай фізично існує лише одне випробування. Беремо довільну подію A , що є підмножиною Ω і повторюємо це випробування n разів, фіксуючи їх результат. В кожному повторі випробування може статися A або $\bar{A}$.

В кожному повторі випробування випадкові фактори якось реалізуються і на це не впливає як вони реалізовувались в попередніх повторах випробувань. З цього випливає наступний прийом математичної статистики.

Результати n повторів фізично існуючого випробування можна вважати результатом одного випробування над n незалежними віртуальними копіями фізично існуючого випробування (результати n підкидань монети однією людиною, можна вважати, що n незалежних копій людини підкинули монету по одному разу). Тоді з кожним i -им віртуальним випробуванням зв'язуємо випадкову величину X_i (віртуальну), що приймає значення 1 або 0 (настала подія A або $\bar{A}$ в i -ому випробуванні). Отже, отримали віртуальну схему n випробувань Бернуллі, для якої вірна теорема Бернуллі, а це означає, що будь-які реалізації випадкової величини $\frac{S_n}{n}$ при достатньо великому n практично не відрізняються від p .

Але будь-яка реалізація $\frac{S_n}{n}$ є частістю настання події A в довільних n повторів фізично існуючого випробування - $W_n(A)$.

Таким чином, $W_n(A)$ в довільній необмеженій серії випробувань за ймовірністю збігається до аксіоматичної за Колмогоровим ймовірності настання події A .

ЛЕКЦІЯ 34

Закон великих чисел

Маємо послідовність незалежних і однаково розподілених випадкових величин $X_1 \dots X_n$.

$$\forall M X_i = \nu, |\nu| < \infty, D X_i = \sigma^2 < \infty$$

Примітка.

Якщо дві випадкові величини мають однакову функцію розподілу або щільності, то їх моменти однакові.

$\overline{X_n}$ – середнє арифметичне цих випадкових величин.

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$M \overline{X_n} = \nu$$

$$D \overline{X_n} = \frac{\sigma^2}{n}$$

$$0 \leq P(|\overline{X_n} - \nu| > \varepsilon) \leq \frac{M(\overline{X_n} - \nu)^2}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

$$\exists \forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|\overline{X_n} - \nu| > \varepsilon) = 0$$

Послідовність середньоарифметичних незалежних однаково розподілених випадкових величин збігається по ймовірності до математичного сподівання кожної з них (до ν).

Наслідок 1

Теорема Бернуллі є частковим випадком Закону великих чисел, де $\nu = p, \sigma^2 = \frac{pq}{n}$.

Наслідок 2

Якщо над однією випадковою величиною повторено n раз випробування, де n достатньо велике, то середнє арифметичне результатів випробування є «хорошою» оцінкою математичного сподівання цієї випадкової величини.

Центральна гранична теорема

Маємо послідовність незалежних і однаково розподілених випадкових величин $X_1 \dots X_n$: $|\nu| < \infty, \sigma^2 < \infty$

$$S_n = \sum_{i=1}^n X_i,$$

тоді **нормована випадкова величина** $\overline{S}_n$ має вигляд:

$$\overline{S}_n = \frac{\sum_{i=1}^n (X_i - \nu)}{\sigma\sqrt{n}}, M\overline{S}_n = 0, D\overline{S}_n = 1.$$

Формулювання теореми:

Послідовність незалежних нормальних сум збігається до випадкової величини X , що має нормований нормальний розподіл $n(x, 0, 1)$.

$$n(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1(x-0)^2}{2 \cdot 1}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (n(x, \nu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(x-\nu)^2}{2 \sigma^2}}).$$

У граничній випадкової величини математичне сподівання 0 і дисперсія 1, тому, що у $\overline{S}_n$ для будь-якого n математичне сподівання дорівнює 0 і дисперсія дорівнює 1.

Примітка. Далі буде показано, що в теоремі використовується четверте означення збіжності – збіжність по розподілу.

Наслідки

- 1) При доведенні цієї теореми не обов'язково нормувати $\overline{S}_n$. Це робиться для того, щоб граничний нормальний розподіл не мав нескінчених параметрів.
- 2) З цього випливає наступне: як показує досвід, центральна гранична теорема реалізована при $n \geq 40$. Тоді:

3) $\sum_{i=1}^n X_i$ наближено має нормальний розподіл $n(x, nv, \sqrt{n\sigma^2})$

4) $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ наближено має нормальний розподіл $n(x, v, \frac{\sigma}{\sqrt{n}})$

5) **Теорема Муавра-Лапласа.** Для достатньо великого n ($n \geq 40$), випадкова величина, що має нормований біноміальний розподіл, наближено розподілена по нормованому нормальному закону.

Дійсно, розглянемо випадкову величину, що має біноміальний розподіл. Її сума - $\sum_{i=1}^n X_i$, X_i взято з теореми Бернуллі (незалежні, однаково розподілені, $M(X_i) = p$, $D(X_i) = pq$).

Тоді граничний розподіл для $\bar{S}_n = \frac{\sum_{i=1}^n (X_i - p)}{\sqrt{npq}} \in n(x, 0, 1)$.

Наслідок: Нехай над довільною випадковою величиною повторюються випробування n разів, тоді сума результатів випробування є реалізація випадкової величини, що має нормальний розподіл з параметрами $n(x, nv, \sqrt{n\sigma^2})$. А реалізація середнього арифметичного – це реалізація випадкової величини з параметрами $n(x, v, \frac{\sigma}{\sqrt{n}})$.

Загальний випадок центральної граничної теореми (дається без доведення і лише її інженерна інтерпретація): Сума достатньо великої кількості ($n \rightarrow \infty$) незалежних, будь-як розподілених випадкових величин кожна з них не домінує над іншими (випробування приймають значення одного порядку), має нормальний розподіл.

Саме використовуючи загальний випадок центральної граничної теореми, в теорії похибок вимірювання похибка вимірювання вважається розподіленою нормально.

*****Теорема про еквівалентне означення збіжності з ймовірністю 1:**

$$P\left(\omega: \bigcap_{r=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} (\varphi_n(\omega) - \varphi(\omega)) \leq \frac{1}{r}\right) = 1$$

Покажемо, що це означення еквівалентне наступному:

$$\forall r > 0 \quad \exists \lim_{m \rightarrow \infty} P\left(\omega: \sup_{n \geq m} |\varphi_n(\omega) - \varphi(\omega)| > \frac{1}{r}\right) = 0$$

Примітка. Еквівалентне означення збіжності з ймовірністю одиниця можна використовувати також у вигляді:

$$\forall r > 0, \quad \lim_{m \rightarrow \infty} P(\sup_{n \geq m} |X_n - X| > \frac{1}{r}) = 0$$

Наслідок: Збіжність з ймовірністю 1 швидше за збіжність по ймовірності, оскільки при $m = n$ $\sup$ зникає.

Доведення теореми

Побудуємо подію

$$B = \bar{A},$$

$$A = \left\{ w : \bigcap_{r=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} |\varphi_n(w) - \varphi(w)| \leq \frac{1}{r} \right\}, P(A)=1 \Rightarrow$$

$$B = \left\{ w : \bigcup_{r=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} |\varphi_n(w) - \varphi(w)| > \frac{1}{r} \right\}, P(B) = 0.$$

ЛЕКЦІЯ 35

$$P\left(\omega : \bigcap_r \bigcup_m \bigcap_{n \geq m} \left(|\varphi_n(\omega) - \varphi(\omega)| \leq \frac{1}{r} \right)\right) = 1,$$

$$\forall r > 0 \quad \exists \lim_{n \rightarrow \infty} P\left(\omega : \sup_{n \geq m} |\varphi_n(\omega) - \varphi(\omega)| > \frac{1}{r}\right) = 0.$$

$$B = \left\{ \bigcup_{r=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} |\varphi_n(w) - \varphi(w)| > \frac{1}{r} \right\}$$

Задамо подію B_{rm} :

$$B_{rm} = \left\{ \omega : \bigcup_{n \geq m} \left(|\varphi_n(\omega) - \varphi(\omega)| > \frac{1}{r} \right) \right\} = \left\{ \omega : \sup_{n \geq m} |\varphi_n(\omega) - \varphi(\omega)| > \frac{1}{r} \right\},$$

де $r = \text{const}$, $m = \text{const}$.

Примітка. Довести самостійно, взявши довільне ω з правої частини рівності, покажемо, що воно належить події з лівої частини рівності, і навпаки.

$$B_r = \bigcap_{m=1}^{\infty} B_{rm}, \quad r = \text{const}.$$

За побудовою $B_{r(m+1)} \subset B_{rm}$.

Покажемо, що $P(B_r) = \lim_{m \rightarrow \infty} P(B_{rm})$ (ця рівність була без доведення використана в попередніх лекціях).

Доведення.

$$P(B_r) = 1 - P(\bar{B}_r) = 1 - P\left(\bigcup_{m=1}^{\infty} \bar{B}_{rm}\right)$$

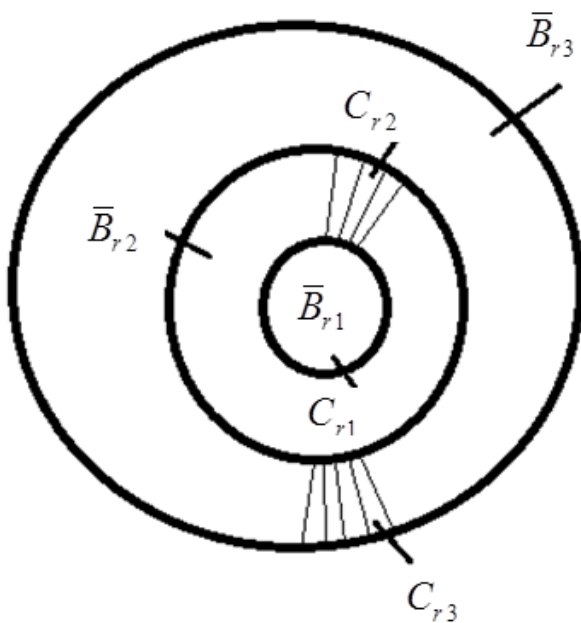
$$\bar{B}_{rm} \subset \bar{B}_{r(m+1)}$$

$$C_{r1} = \bar{B}_{r1}, C_{r2} = \bar{B}_{r2} / \bar{B}_{r1}, C_{r3} = \bar{B}_{r3} / \bar{B}_{r2}, \dots,$$

$$C_{rm} = \bar{B}_{rm} / \bar{B}_{r(m-1)}$$

$$\forall C_{ri} \cap C_{rj} = \emptyset$$

За побудовою:



$$\bigcup_{j=1}^m \bar{B}_{rj} = \bigcup_{j=1}^m C_{rj}$$

$$\bar{B}_{rm} = \bigcup_{j=1}^m C_{rj}$$

$$\bigcup_{m=1}^{\infty} \bar{B}_{r m} = \bigcup_{m=1}^{\infty} C_{r m}$$

Отже, матимемо

$$\begin{aligned} P(B_r) &= 1 - P\left(\bigcup_{m=1}^{\infty} \bar{B}_{r m}\right) = 1 - P\left(\bigcup_{m=1}^{\infty} C_{r m}\right) = \left| \begin{array}{l} \text{III аксіома} \\ \text{Колмогорова} \end{array} \right| = 1 - \sum_{m=1}^{\infty} P(C_{r m}) = \\ &= \lim_{m \rightarrow \infty} \left(1 - \sum_{j=1}^m P(C_{r j})\right) = \lim_{m \rightarrow \infty} \left(1 - P\left(\bigcup_{j=1}^m C_{r j}\right)\right) = \left| \begin{array}{l} \text{III аксіома Колмогорова} \end{array} \right| = \\ &= \lim_{m \rightarrow \infty} \left(1 - P(\bar{B}_{r m})\right) = \lim_{m \rightarrow \infty} P(B_{r m}). \end{aligned}$$

$$P(B_{r m}) = \lim_{m \rightarrow \infty} P\left(\omega: \sup_{n \geq m} |\varphi_n(\omega) - \varphi(\omega)| > \frac{1}{r}\right).$$

Залишилось показати, що $\lim_{m \rightarrow \infty} P(B_{r m}) = 0 \quad \forall r \rightarrow \infty$.

$$B = \bigcup_{r=1}^{\infty} B_r, \quad B_r \subset B_{r+1}.$$

Звідки для того, щоб $P(B) = 0$ потрібно, щоб для будь-якого натурального r $P(B_r) = 0$.

***Посилений закон великих чисел

Примітка. При доведенні посиленого закону великих чисел, збіжність з ймовірністю 1 буде використана в наступному вигляді:

$$\forall r > 0 \quad \lim_{m \rightarrow \infty} P\left(\omega: \sup_{n \geq m} |X_n - X| > \frac{1}{r}\right) = 0.$$

Нехай $X_1, X_2, \dots, X_n, \dots$ – послідовність незалежних довільно розподілених випадкових величин.

При цьому виконується

$$\forall n \quad |MX_n| < \infty, \quad \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty.$$

Переходимо до нового ряду незалежних випадкових величин:

$$X_1 - v_{X_1}, X_2 - v_{X_2}, \dots, X_n - v_{X_n}, \dots$$

який для спрощення позначатимемо як $X_1, X_2, \dots, X_n, \dots$

$$\forall n \quad MX_n = 0, \text{ а для дисперсії виконується } \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty.$$

Примітка. Виконується властивість $D(X + C) = DX$.

Формулювання посиленого закону великих чисел:

$$\lim_{m \rightarrow \infty} P \left(\sup_{n \geq m} \left| \frac{1}{n} \sum_{i=1}^n X_i \right| > \varepsilon \right) = 0$$

$$\forall \varepsilon > 0 \quad (\varepsilon \text{ раніше було } \frac{1}{r}, \text{ границя } h = 0).$$

Покажемо, що закон великих чисел і теорема Бернуллі є частковим випадком посиленого закону великих чисел.

Маємо послідовність незалежних однаково розподілених випадкових величин $X_1, X_2, \dots, X_n, \dots$ $|v| < \infty, \sigma^2 < \infty$.

$$\sigma^2 \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

$$X_1 - v, X_2 - v, \dots, X_n - v, \dots$$

Тоді має місце посилений закон великих чисел:

$$\lim_{m \rightarrow \infty} P \left(\sup_{n \geq m} \left| \sum_{i=1}^n \frac{(X_i - v)}{n} \right| > \varepsilon \right) = 0$$

$$\lim_{m \rightarrow \infty} P \left(\sup_{n \geq m} \left| \frac{\sum_{i=1}^n X_i}{n} - v \right| > \varepsilon \right) = 0, \quad \forall \varepsilon > 0$$

А оскільки теорема Бернуллі – частковий випадок закону великих чисел: $MX_i = p$, $DX_i = pq$.

Таким чином, частість настання події A в довільних повторних випробуваннях з ймовірністю 1 збігається до $P(A)$ (ймовірність настання A за Колмогоровим).

Доведення посиленого закону великих чисел.

Лема Колмогорова (аналог нерівності Чебишова).

$$P\left(\max_{1 \leq m \leq n} \left| \sum_{i=1}^m X_i \right| > \varepsilon\right) \leq \frac{D\left(\sum_{i=1}^n X_i\right)}{\varepsilon^2} = \frac{D(S_n)}{\varepsilon^2} = \frac{M(S_n^2)}{\varepsilon^2},$$

$$\forall \varepsilon > 0, \quad S_m = \sum_{i=1}^m X_i, \quad S_n = \sum_{i=1}^n X_i, \quad M(S_n) = 0.$$

ЛЕКЦІЯ 36

Доведення.

Розглянемо подію $A = (\vartheta = m)$, яка настає тільки тоді, коли виконується наступні нерівності:

$$|X_1| \leq \varepsilon, \left| \sum_{i=1}^2 X_i \right| \leq \varepsilon, \dots, \left| \sum_{i=1}^{m-1} X_i \right| \leq \varepsilon, \left| \sum_{i=1}^m X_i \right| > \varepsilon.$$

$$(X_1, X_2, \dots, X_m) \in A = \{\forall x \in R_m, |x_1| \leq \varepsilon, \sum_{i=1}^2 x_i \leq \varepsilon, \dots, \sum_{i=1}^{m-1} x_i \leq \varepsilon, \sum_{i=1}^m x_i > \varepsilon\}.$$

Розглянемо наступну подію $(X_1, X_2, \dots, X_m) \in A \subset R_m$.

Індикаторною функцією $1_{X_1, X_2, \dots, X_m \in A}(X_1, X_2, \dots, X_m) = 1_A(X_1, X_2, \dots, X_m)$

називається числова, скалярна функція m випадкових аргументів, яка приймає значення 1, якщо внаслідок випробування $(X_1, X_2, \dots, X_m)$ прийме значення, що належить A , і 0 – в протилежному випадку.

Властивості індикаторної функції:

$$1) M1_A(X_1, X_2, \dots, X_m) = P(X_1, X_2, \dots, X_m \in A).$$

Доведемо для цей факт для неперервного випадку, для дискретного – самостійно.

За означенням

$$M\varphi(X_1, X_2, \dots, X_m) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \varphi(x_1, x_2, \dots, x_m) f(x_1, x_2, \dots, x_m) dx_1 \dots dx_m$$

$$M1_{\vartheta=m}(X_1, X_2, \dots, X_m)$$

$$= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} 1_{X_1, X_2, \dots, X_m \in A}(x_1, x_2, \dots, x_m) f(x_1, x_2, \dots, x_m) dx_1 \dots dx_m =$$

$$= P(X_1, X_2, \dots, X_m \in A).$$

Примітка: m –кратний інтеграл по нескінченим границям дорівнює m –вимірному інтегралу по арифметичному простору R_m .

$$2) M1_{\vartheta=m} = P(\vartheta = m).$$

$$3) \bigcup_{m=1}^n (\vartheta = m) = \max_{1 \leq m \leq n} \{ |S_m| > \varepsilon, \forall \varepsilon > 0 \}.$$

Дійсно, якщо виконується подія справа, то виконується подія зліва і навпаки.

Події $\{\vartheta = m\}$ для різних m – попарно несумісні, що випливає з означення. Тобто якщо декілька модулів більше ε , то виконується лише одна подія $\{\vartheta = m\}$, у якій m – мінімальний серед усіх модулів, кожен з яких більше ε .

$$4) X \geq Y \Rightarrow MX \geq MY.$$

$$X - Y \geq 0 \Rightarrow M(X - Y) \geq 0 \Rightarrow MX \geq MY.$$

5) Якщо $X_1, X_2, \dots, X_n$ – незалежні, то будь-які k з них – незалежні.

Доведемо для неперервного випадку.

$$f(x_1, x_2, \dots, x_m) = \prod_{i=1}^n f_{x_i}(x_i).$$

$$f_{x_1, x_2, \dots, x_{i_k}}(x_1, x_2, \dots, x_{i_k}) = \prod_{l=1}^k f_{x_{i_l}}(x_{i_l}) \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \prod_{t=1}^{n-k} f_{x_{j_t}}(x_{j_t}) dx_{j_1} \dots dx_{j_{n-k}}$$

Очевидно, що n -к-кратний інтеграл дорівнює одиниці, тому, що

$$\left(\int_{-\infty}^{+\infty} f_{x_{j_t}}(x_{j_t}) dx_{j_t} = 1 \right).$$

6) $X_1, X_2, \dots, X_n$ – незалежні,

$$\varphi(x_1, x_2, \dots, x_n) = \varphi_1(x_1, x_2, \dots, x_m) \varphi_2(x_{m+1}, x_{m+2}, \dots, x_n),$$

де множники - неперервні або мають обмежену кількість стрибків першого роду, то

$$M\varphi(X_1, X_2, \dots, X_n) = M(\varphi_1(X_1, X_2, \dots, X_m) \varphi_2(X_{m+1}, X_{m+2}, \dots, X_n)) =$$

$$M\varphi_1(X_1, X_2, \dots, X_m) M\varphi_2(X_{m+1}, X_{m+2}, \dots, X_n).$$

Доведення:

$$\begin{aligned} & M(\varphi_1(X_1, X_2, \dots, X_m) \varphi_2(X_{m+1}, X_{m+2}, \dots, X_n)) = \\ & \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \varphi_1(x_1, x_2, \dots, x_m) \varphi_2(x_{m+1}, x_{m+2}, \dots, x_n) \prod_{i=1}^m f_{x_i}(x_i) \prod_{j=m+1}^n f_{x_j}(x_j) dx_1 \dots dx_n \\ & = M\varphi_1(X_1, X_2, \dots, X_m) M\varphi_2(X_{m+1}, X_{m+2}, \dots, X_n). \end{aligned}$$

7) $X_1, X_2, \dots, X_n$ – незалежні, то

$$M\left(\prod_{i=1}^n \varphi_i(X_i)\right) = \prod_{i=1}^n M\varphi_i(X_i),$$

Де $\varphi_i(X_i)$ – числові скалярні функції одного аргументу (неперервні бо мають обмежену кількість стрибків першого роду). Доведення аналогічне.

$$8) \quad M(S_m \cdot 1_{\vartheta=m} \cdot X_k) = 0, k > m, \quad S_m = \sum_{i=1}^m X_i, \quad S_m \cdot 1_{\vartheta=m} = \varphi_1(X_1, X_2, \dots, X_m), \\ \varphi_2(X_k) = X_k, \text{ послідовність незалежних випадкових величин } X_1, X_2, \dots, X_m, X_k.$$

Далі використовуємо пункт 6.

$$9) \quad M(S_n^2 1_{\vartheta=m}) \geq M(S_m^2 1_{\vartheta=m}).$$

$$M(S_n^2 1_{\vartheta=m}) = M((S_m + S_{n-m})^2 1_{\vartheta=m}) =$$

Розкриваємо квадрат і обчислюємо математичне сподівання:

$$= M(S_m^2 1_{\vartheta=m}) + 2M(S_m S_{n-m} 1_{\vartheta=m}) + M(S_{n-m}^2 1_{\vartheta=m}) \quad (S_{n-m} = \sum_{i=m+1}^n X_i).$$

Другий доданок дорівнює нулю, а третій більший або рівний нулю.

10) $S_n^2 \geq S_n^2 \sum_{m=1}^n 1_{\vartheta=m}$. (Оскільки події $(\vartheta = m)$ – попарно несумісні в цій сумі лише один доданок може бути одиницею, інші нулі, чи усі доданки нулі.

$$11) \quad DS_n = MS_n^2 \quad (MS_n = 0, MX_i = 0 \quad \forall i = \overline{1, n}).$$

$$DS_n = MS_n^2 \geq MS_n^2 \sum_{m=1}^n 1_{\vartheta=m}.$$

ЛЕКЦІЯ 37

$$DS_n = MS_n^2 \geq M\left(S_n^2 \sum_{m=1}^n 1_{(\vartheta=m)}\right) = M\sum_{m=1}^n S_n^2 \cdot 1_{(\vartheta=m)} = \\ = \sum_{m=1}^n M\left(S_n^2 \cdot 1_{(\vartheta=m)}\right) \geq \sum_{m=1}^n M\left(S_m^2 \cdot 1_{(\vartheta=m)}\right) = M\left(\sum_{m=1}^n S_m^2 \cdot 1_{(\vartheta=m)}\right)$$

Оскільки

$$\sum_{m=1}^n S_m^2 \cdot 1_{(\vartheta=m)} \geq \varepsilon^2 \sum_{m=1}^n 1_{(\vartheta=m)},$$

тому, що в правій сумі можуть бути всі доданки нулі, або лише одна одиниця, при якій відповідний модуль більше ε , а інші доданки в сумі – нулі.

Тоді

$$M\left(\sum_{m=1}^n S_m^2 \cdot 1_{(\vartheta=m)}\right) \geq M\left(\varepsilon^2 \sum_{m=1}^n 1_{(\vartheta=m)}\right) \quad (\text{властивість 4}).$$

$$\begin{aligned}
M\left(\sum_{m=1}^n S_m^2 \cdot 1_{(v=m)}\right) &\geq \varepsilon^2 M\left(\sum_{m=1}^n 1_{(v=m)}\right) = \varepsilon^2 \sum_{m=1}^n M(1_{(v=m)}) = \varepsilon^2 \sum_{m=1}^n P(v=m) = \\
&= \varepsilon^2 P\left(\bigcup_{m=1}^n (v=m)\right) = \varepsilon^2 P\left(\max_{1 \leq m \leq n} |S_m| > \varepsilon\right),
\end{aligned}$$

Примітка. Показати самостійно, що

$$\bigcup_{m=1}^n (v=m) = \max_{1 \leq m \leq n} |S_m| > \varepsilon.$$

Таким чином $\forall \varepsilon > 0$

$$DS_n \geq \varepsilon^2 P\left(\max_{1 \leq m \leq n} |S_m| > \varepsilon\right).$$

Нерівність Колмогорова доведена.

Доведемо посилений закон великих чисел,

$$\exists \lim_{m \rightarrow \infty} P\left(\sup_{n \geq m} \left|\frac{S_n}{n}\right| > \varepsilon\right) = 0,$$

при додаткових умовах:

$$MX_i = 0, \quad S_n = \sum_{i=1}^n X_i, \quad \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty.$$

$$A_k = \left\{ \max_{2^{k-1} \leq n \leq 2^k} \left| \frac{S_n}{n} \right| > \varepsilon \right\}$$

$$\begin{aligned}
P(A_k) &= P\left(\max_{2^{k-1} \leq n \leq 2^k} \left| \frac{S_n}{n} \right| > \varepsilon\right) \leq P\left(\max_{2^{k-1} \leq n \leq 2^k} |S_n| > \varepsilon \cdot 2^{k-1}\right) \leq \\
&\leq P\left(\max_{1 \leq n \leq 2^k} |S_n| > \varepsilon \cdot 2^{k-1}\right) \leq \left| \begin{array}{l} \text{використовуємо} \\ \text{лему Колмогорова} \end{array} \right| \leq \frac{DS_{2^k}}{\varepsilon^2 \cdot 2^{2(k-1)}} = \varepsilon^{-2} \cdot 2^{-2k} \cdot 4 \sum_{n \leq 2^k} \sigma_n^2. \\
\sum_{k=1}^{\infty} P(A_k) &\leq \varepsilon^{-2} \cdot 4 \sum_{k=1}^{\infty} 2^{-2k} \sum_{n \leq 2^k} \sigma_n^2.
\end{aligned}$$

Змінимо порядок сумування. Розглянемо низку нерівностей: $n \leq 2^k$.

При фіксованому n ($n \leq 2^k \Rightarrow n^2 \leq 2^{2k} \Rightarrow 2^{-2k} \leq n^{-2}$), беремо мінімально можливе k , для якого ще виконуються $2^{-2k} \leq n^{-2}$, і від цього значення k спрямовуємо до нескінченності. Тоді

$$\begin{aligned} \sum_{k=1}^{\infty} P(A_k) &\leq \varepsilon^{-2} \cdot 4 \sum_{k=1}^{\infty} 2^{-2k} \sum_{n \leq 2^k} \sigma_n^2 = \varepsilon^{-2} \cdot 4 \sum_{n=1}^{\infty} \sigma_n^2 \sum_{2^{-2k} \leq n^{-2}} 2^{-2k} \leq \\ &\leq \varepsilon^{-2} \cdot 4 \sum_{n=1}^{\infty} \sigma_n^2 \cdot \left(n^{-2} + \frac{n^{-2}}{4} + \frac{n^{-2}}{16} + \dots \right) = \varepsilon^{-2} \cdot 4 \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} \cdot \frac{1}{1 - \frac{1}{4}} < \infty \\ &\Rightarrow \sum_{k=1}^{\infty} P(A_k) < \infty. \end{aligned}$$

Примітка. В круглих дужках для другої суми замість її доданків записані їх верхні оцінки і n^{-2} виносимо за дужки, а сума нескінченно спадної геометричної, що залишилась в круглих дужках, дорівнює $\frac{1}{1 - \frac{1}{4}}$.

$$P_{2^{m-1}} = P\left(\sup_{n \geq 2^{m-1}} \left| \frac{S_n}{n} \right| > \varepsilon\right) = P\left(\bigcup_{k \geq m} A_k\right) \leq \sum_{k=m}^{\infty} P(A_k) \Rightarrow$$

Примітка.

$$\begin{aligned} A_k &= \max_{2^{k-1} \leq n \leq 2^k} \left| \frac{S_n}{n} \right| > \varepsilon \\ \left\{ \sup_{n \geq 2^{m-1}} \left| \frac{S_n}{n} \right| > \varepsilon \right\} &= \left\{ \bigcup_{k \geq m} A_k \right\} \end{aligned}$$

Оскільки

$$\sum_{k=1}^{\infty} P(A_k) < \infty,$$

З цієї нерівності випливає, що

$$\exists \lim_{m \rightarrow \infty} \sum_{k=m}^{\infty} P(A_k) = 0,$$

таким чином, довели що існує ліміт

$$\lim_{m \rightarrow \infty} P_{2^{m-1}} = P\left(\sup_{n \geq 2^{m-1}} \left| \frac{S_n}{n} \right| > \varepsilon\right) = 0,$$

чи існує границя числової послідовності

$$\lim_{m \rightarrow \infty} P_{2^0}, P_{2^1}, P_{2^2}, \dots, P_{2^{m-1}}, \dots = 0$$

а нам треба

$$\lim_{m \rightarrow \infty} P_1, P_2, \dots, P_m, \dots = 0$$

Але ця числова послідовність обмежена і монотонно не зростаюча, тому має завжди границю і будь-яка її нескінченна підпослідовність має ту ж саму границю. Закон великих чисел доведено.

ЛЕКЦІЯ 38

***Комплексні випадкові величини

Комплексною випадковою величиною Z будемо називати комплексно значну функцію двох дійсних випадкових аргументів

$$Z = X + iY$$

де X, Y – дійсні випадкові величини, а i – уявна одиниця, $i^2 = -1$.

Примітка. Комплексні випадкові величини вводяться для вивчення властивостей характеристичної функції $\varphi_X(t)$. Дійсно, за означенням

$$\varphi_X(t) = M e^{iXt}$$

А за формулою Ейлера $e^{iu} = \cos u + i \sin u$

$$\varphi_X(t) = M(\cos Xt + i \sin Xt).$$

В дужках комплексна випадкова величина.

Операції над комплексними випадковими величинами такі самі, що і над звичайними комплексними числами, з тією різницею, що тепер дійсна частина та коефіцієнт при уявній частині результату є дійсними випадковими величинами.

Так **сума двох комплексних випадкових величин** $Z_1 = X_1 + iY_1$ та $Z_2 = X_2 + iY_2$ визначається як

$$Z_1 + Z_2 = X_1 + X_2 + i(Y_1 + Y_2)$$

Добуток цих величин є

$$\begin{aligned} Z_1 \cdot Z_2 &= (X_1 + iY_1) \cdot (X_2 + iY_2) = X_1X_2 + iX_1Y_2 + iX_2Y_1 - Y_1Y_2 \\ &= X_1X_2 - Y_1Y_2 + i(X_1Y_2 + X_2Y_1). \end{aligned}$$

Мають місце наступні формули:

$$M\varphi(x_1, \dots, x_n, i) = \sum_{j_1=1}^{s_1} \dots \sum_{j_n=1}^{s_n} \varphi(x_{j_1}^1, \dots, x_{j_n}^1, i) p(x_{j_1}^1, \dots, x_{j_n}^{n1}) -$$

дискретний випадок.

$$M\varphi(x_1, \dots, x_n, i) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \varphi(x_1, \dots, x_n, i) f(x_1, \dots, x_n) dx_1 \dots dx_n -$$

неперервний випадок.

$$M \sum_{l=1}^k \varphi_l(x_{l_1} \dots x_{l_{k_l}}, i) = \sum_{l=1}^k M(\varphi_l(x_{l_1} \dots x_{l_{k_l}}, i)).$$

Доведення цієї формули для дискретного і неперервного випадків очевидно, бо для операції математичного сподівання уявна одиниця – це константа.

З цих формул випливає

$$M(iY) = iM(Y), \text{ бо } i - \text{це константа.}$$

$$M(X + iY) = MX + iMY.$$

$$\begin{aligned} M(Z_1 \cdot Z_2) &= M(X_1 + iY_1)(X_2 + iY_2) \\ &= M(X_1 \cdot X_2) - M(Y_1 \cdot Y_2) + i(M(X_1 \cdot Y_2) - M(X_2 \cdot Y_1)). \end{aligned}$$

$$M \sum_{j=1}^n Z_j = M \sum_{j=1}^n (X_j + iY_j) = \sum_{j=1}^n M(X_j + iY_j) = \sum_{j=1}^n MZ_j.$$

Нехай є випадкові величини X_1, X_2, Y_1, Y_2 . Пари $X_1X_2, X_1Y_2, X_2Y_1, Y_1Y_2$ є незалежними.

Тоді, $Z_1 = X_1 + iY_1$ та $Z_2 = X_2 + iY_2$ називаються **незалежними**.

Якщо Z_1, Z_2 – незалежні, то математичне сподівання їх добутку дорівнює добутку їх математичних сподівань.

Доведення:

$$\begin{aligned} M(Z_1 \cdot Z_2) &= M((X_1 + iY_1) \cdot (X_2 + iY_2)) \\ &= M(X_1 \cdot X_2) - M(Y_1 \cdot Y_2) + i(M(X_1 \cdot Y_2) - M(X_2 \cdot Y_1)) \\ &= MX_1 \cdot MX_2 - MY_1 \cdot MY_2 + i(MX_1 \cdot MY_2 - MX_2 \cdot MY_1) \\ &= (MX_1 + iMY_1) \cdot (MX_2 + iMY_2) = MZ_1 \cdot MZ_2. \end{aligned}$$

Розглянемо n – випадкових величин $Z_j = X_j + iY_j, j = \overline{1, n}$.

2^n n – вимірних випадкових величин

$(X_1 \vee Y_1), (X_2 \vee Y_2), \dots, (X_n \vee Y_n)$ – незалежні.

Тоді комплексні випадкові величини Z_j – незалежні.

Довести самостійно, що математичне сподівання добутку випадкових величин дорівнює добутку їх математичних сподівань:

$$M \prod_{j=1}^n Z_j = \prod_{j=1}^n M Z_j.$$

Модуль комплексної випадкової величини визначається як

$$|Z| = \sqrt{X^2 + Y^2}.$$

Доведемо важливу нерівність для комплексної випадкової величини (для дійсної випадкової величини ми вже робили це раніше):

$$|MZ| \leq M |Z|,$$

Розглянемо лише два випадки:

1. Нехай (X, Y) – дискретна випадкова величина, компоненти якої приймають скінченну кількість значень: $x_l, l = \overline{1, s}$ та $y_j, j = \overline{1, m}$, відповідно. Задано їх сумісний розподіл $p(x_l, y_j) = P(X = x_l, Y = y_j), l = \overline{1, s}, j = \overline{1, m}$.

Тоді

$$|MZ| = \left| \sum_{l=1}^s \sum_{j=1}^m (x_l + i y_j) p(x_l, y_j) \right| \leq \sum_{l=1}^s \sum_{j=1}^m |x_l + i y_j| p(x_l, y_j) = \sum_{l=1}^s \sum_{j=1}^m \sqrt{x_l^2 + y_j^2} p(x_l, y_j) = M |Z|$$

У дискретному випадку нерівність доведено.

2. Неперервний випадок. Нехай (X, Y) – неперервна випадкова величина зі щільністю розподілу $f(x, y)$. Тоді

$$|MZ| = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + i y) f(x, y) dx dy \right| \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x + i y| f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sqrt{x^2 + y^2} f(x, y) dx dy = M |Z|.$$

***Характеристичні функції. Властивості

Як було показано вище, *характеристичною функцією* дійсної випадкової величини X називається функція $\phi_X(t)$, що дорівнює

$$\phi_X(t) = M e^{itX}, \quad (38.1)$$

де t – дійсний параметр.

Якщо розкрити в (38.1) e^{itX} за формулою Ейлера $e^{itX} = \cos tX + i \sin tX$, то одержуємо

$$\phi_X(t) = M \cos tX + iM \sin tX. \quad (38.2)$$

Для дискретної випадкової величини X , що приймає значення $x_j, j = \overline{1, s}$, з ймовірностями $p(x_j), j = \overline{1, s}$, характеристична функція має вигляд

$$\phi_X(t) = \sum_{j=1}^s e^{ix_j t} p(x_j). \quad (38.3)$$

У випадку неперервної випадкової величини X зі щільністю розподілу $f_X(x)$

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{ixt} f_X(x) dx, \quad (38.4)$$

тобто $\phi_X(t)$ є ніщо інше, як перетворення Фур'є щільності $f_X(t)$.

Розглянемо властивості характеристичних функцій.

1. Характеристична функція обмежена.

$$|\phi_X(t)| \leq 1, \phi_X(0) = 1.$$

Дійсно, з нерівності $|MZ| \leq M|Z|$ маємо

$$|Me^{itX}| \leq M|e^{itX}| = M|\cos tX + i \sin tX| = M(\cos^2 tX + \sin^2 tX) = 1.$$

Це, зокрема, означає, що $\phi_X(t)$ існує для довільних t .

2. Нехай $Y = aX + b$, де a та b – сталі.

Тоді

$$\varphi_Y(t) = M(e^{i(aX+b)t}) = e^{itb} M(e^{itaX}) = e^{itb} \varphi_X(at). \quad (38.5)$$

ЛЕКЦІЯ 39

Розглянемо низку допоміжних результатів.

Перший додатковий результат

Нехай $X_1, X_2, \dots, X_n$ – незалежні. Розглянемо випадкові величини $\varphi_1(X_1, \dots, X_k), \varphi_2(X_{k+1}, \dots, X_n)$.

$\varphi_1(x_1, \dots, x_k), \varphi_2(x_{k+1}, \dots, x_n)$ – неперервні числові скалярні функції k та $n - k$ аргументів відповідно.

Доведемо, що $\varphi_1(X_1, \dots, X_k), \varphi_2(X_{k+1}, \dots, X_n)$ – незалежні випадкові величини.

Розглянемо дві довільні числові множини A і B : $A \subset R, B \subset R$, тоді має місце рівність :

$$P\left(\begin{array}{l} \varphi_1(X_1 \dots X_k) \in A \\ \varphi_2(X_{k+1} \dots X_n) \in B \end{array}\right) = P((\varphi_1(X_1 \dots X_k) \in A) \cdot P(\varphi_2(X_{k+1} \dots X_n) \in B)).$$

Вводимо наступні індикаторні функції

$$1_{\varphi_1(X_1 \dots X_k) \in A}(X_1 \dots X_k) = 1_{\varphi_1} = \begin{cases} 1, & \varphi_1(X_1 \dots X_k) \in A \\ 0, & \text{в протилежному випадку} \end{cases}$$

$$1_{\varphi_2(X_{k+1} \dots X_n) \in B}(X_{k+1} \dots X_n) = 1_{\varphi_2} = \begin{cases} 1, & \varphi_2(X_{k+1} \dots X_n) \in B \\ 0, & \text{в протилежному випадку} \end{cases}$$

$$1_{\substack{\varphi_1(X_1 \dots X_k) \in A, \\ \varphi_2(X_{k+1} \dots X_n) \in B}}(X_1 \dots X_n) = 1_{\varphi_1 \varphi_2} = \begin{cases} 1, & \begin{array}{l} \varphi_1(X_1 \dots X_k) \in A \\ \varphi_2(X_{k+1} \dots X_n) \in B \end{array} \\ 0, & \text{в протилежному випадку} \end{cases}$$

Тоді за побудовою

$$1_{\varphi_1 \varphi_2} = 1_{\varphi_1} \cdot 1_{\varphi_2}$$

$$M(1_{\varphi_1 \varphi_2}) = M(1_{\varphi_1} \cdot 1_{\varphi_2}) = M1_{\varphi_1} \cdot M1_{\varphi_2}$$

Примітка. Для φ_1 і φ_2 неперервних чи кусково-неперервних числових скалярних функцій $M(\varphi_1 \cdot \varphi_2) = M(\varphi_1) \cdot M(\varphi_2)$.

Використовуючи другий додатковий результат леми Колмогорова, матимемо

$$P\left(\begin{array}{l} \varphi_1(X_1 \dots X_k) \in A \\ \varphi_2(X_{k+1} \dots X_n) \in B \end{array}\right) = P(\varphi_1(X_1 \dots X_k) \in A) \cdot P(\varphi_2(X_{k+1} \dots X_n) \in B).$$

Нехай $X_1, \dots, X_n$ – неперервні випадкові величини (для дискретних довести самостійно). Тоді позначимо $\varphi_1(X_1 \dots X_k) = Y$, $\varphi_2(X_{k+1} \dots X_n) = Z$.

(Y, Z) – двовимірна випадкова величина, $f_Y(y), f_Z(z)$ функції щільності випадкових величин Y та Z . Знайдемо функцію $f(y, z)$.

Тоді нашу рівність можна переписати у вигляді

$$\iint_{A \oplus B} f(y, z) dy dz = \int_A f_Y(y) dy \cdot \int_B f_Z(z) dz = \iint_{A \oplus B} f_Y(y) \cdot f_Z(z) dy dz$$

Так як A і B – довільні, маємо нескінченну незлічену множину таких інтегральних рівностей, а це означає, що рівні не тільки об'єми, а й підінтегральні рівності, якщо підінтегральні функції є неперервними чи кусково-неперервними з обмеженою кількістю стрибків 1-го роду.

Таким чином, двовимірна функція щільності дорівнює добутку одновимірних $f(y, z) = f(y) \cdot f(z)$, тож випадкові величини Y та Z – незалежні.

Якщо випадкові величини $X_1, \dots, X_n$ є дискретними, до доведення, що $\varphi_1(X_1 \dots X_k), \varphi_2(X_{k+1} \dots X_n)$ безпосередньо впливає з формули

$$P\left(\begin{matrix} \varphi_1(X_1 \dots X_k) \in A \\ \varphi_2(X_{k+1} \dots X_n) \in B \end{matrix}\right) = P(\varphi_1(X_1 \dots X_k) \in A) \cdot P(\varphi_2(X_{k+1} \dots X_n) \in B).$$

Дійсно, для цього треба покласти $A = \varphi_1(x_{j_1}^1, \dots, x_{j_k}^k), B = \varphi_2(x_{j_{k+1}}^{k+1}, \dots, x_{j_n}^n)$, де $x_{j_1}^1, \dots, x_{j_k}^k$ – довільна елементарна подія k -вимірної випадкової величини $X_1, \dots, X_k$, а $x_{j_{k+1}}^{k+1}, \dots, x_{j_n}^n$ – довільна елементарна подія $n-k$ -вимірної випадкової величини $X_{k+1}, \dots, X_n$.

Примітка. Якщо для різних елементарних подій $X_1, \dots, X_k$ $\varphi_1(\cdot)$ Приймає однакові значення, то для знаходження $P(A)$ треба сумувати ймовірності настання цих елементарних подій. Аналогічні дії виконуються для $\varphi_2(\cdot)$.

Другий додатковий результат:

$X_i, i = 1, \dots, n$ – незалежні, $\varphi_i(x_i), i = \overline{1, n}$ – довільні неперервні функції, $\varphi_i(X_i)$ – довільні випадкові величини. Покажемо, що $\varphi_i(x_i)$ та $\varphi_i(X_i)$ – незалежні.

Вводимо $A_1, \dots, A_n$ – довільні числові множини (довільні відрізки числової осі).

$$1_{\varphi_i(X_i) \in A_i}(X_i) = 1_{\varphi_i} = \begin{cases} 1, \varphi_i(x_i) \in A_i \\ 0, \text{в протилежному випадку} \end{cases} \quad (i = \overline{1, n})$$

$$1_{\varphi_i(X_i) \in A_i, \dots, \varphi_n(X_n) \in A_n}(X_1, \dots, X_n) = 1_{\varphi_1, \dots, \varphi_n} = \begin{cases} 1, \forall \varphi_i(x_i) \in A_i \\ 0, \exists \varphi_j(x_j) \notin A_j \end{cases}$$

$$1_{\varphi_1, \dots, \varphi_n} = \prod_{i=1}^n 1_{\varphi_i} \Rightarrow M1_{\varphi_1, \dots, \varphi_n} = \prod_{i=1}^n M1_{\varphi_i}$$

$$P(\varphi_i(X_i) \in A_i) = \prod_{i=1}^n P((\varphi_i(X_i) \in A_i), (i = \overline{1, n}))$$

(використано 6-ий результат леми Колмогорова).

Доведення для неперервного випадку.

$$Y_i = \varphi_i(X_i), \quad f(y_1, \dots, y_n) - ?$$

$$\begin{aligned} \int_{A_1 \oplus \dots \oplus A_n} \dots \int f(y_1, \dots, y_n) dy_1 \dots dy_n &= \prod_{i=1}^n \int_{A_i} f_{Y_i}(y_i) dy_i = \\ &= \int_{A_1 \oplus \dots \oplus A_n} \dots \int \prod_{i=1}^n f_{Y_i}(y_i) dy_1 \dots dy_n. \end{aligned}$$

Оскільки відрізки A_i ($i = \overline{1, n}$) довільні, то з нескінченної кількості інтегральних рівностей випливає рівність підінтегральних виразів, якщо вони є неперервними або кусково-неперервними з обмеженою кількістю розривів I-го роду. Таким чином, n -вимірна функція щільності дорівнює добутку одновимірних.

Примітка. Дискретний випадок доводиться аналогічно виведенню, наведеному в допоміжному результаті один (довести самим, поклавши $A_i = \varphi_i(x_{j_i}^i)$).

$X_1, \dots, X_n$ – дійсні незалежні випадкові величини. Розглянемо наступні комплексні випадкові величини.

За формулою Ейлера: $e^{itX_j} = \cos tX_j + i \sin tX_j$, $j = \overline{1, n}$.

$$(\cos tX_1 \vee \sin tX_1), (\cos tX_2 \vee \sin tX_2), \dots, (\cos tX_n \vee \sin tX_n)$$

Кожна з множини 2^n n -вимірних випадкових величин є незалежними.

Примітка 1. В кожній дужці залишається одна з двох можливих випадкових величин.

Примітка 2. Якщо 2^n n -вимірних випадкових величин – незалежні,

$$(X_1 \vee Y_1), (X_2 \vee Y_2), \dots, (X_n \vee Y_n)$$

то також незалежні 2^n n -вимірних випадкових величин

$$(f_{1,1}(X_1) \vee f_{1,2}(Y_1)), (f_{2,1}(X_2) \vee f_{2,2}(Y_2)), \dots, (f_{n,1}(X_n) \vee f_{n,2}(Y_n)).$$

Таким чином, комплексні випадкові величини e^{itX_j} – незалежні.

$$M\left(\prod_{j=1}^n e^{itX_j}\right) = \prod_{j=1}^n M e^{itX_j}.$$

Математичне сподівання добутку незалежних комплексних випадкових величин дорівнює добутку їх математичних сподівань.

Властивість 3 характеристичної функції

Якщо $X_1, X_2, \dots, X_n$ – незалежні, то

$$\varphi_{\sum_{j=1}^n X_j}(t) = \prod_{j=1}^n \varphi_{X_j}(t).$$

Доведення.

$$M e^{it \sum_{j=1}^n X_j} = M \left(\prod_{j=1}^n e^{itX_j} \right)$$

Використовуючи другий додатковий результат, отримуємо.

$$M\left(\prod_{j=1}^n e^{itX_j}\right) = \prod_{j=1}^n M e^{itX_j} = \prod_{j=1}^n \varphi_{X_j}(t).$$

Властивість 4 характеристичної функції

Нехай початковий момент дійсної випадкової величини $X : |v_n| < \infty$, тоді

$$1) \frac{d^k \varphi_X(t)}{dt^k} = i^k M(X^k \cdot e^{itX}) \Rightarrow \varphi_X^{(k)}(t) |_{t=0} = i^k v_k, \quad k = \overline{1, n}$$

2) Розклад в ряд Маклорена

$$\varphi_X(t) = 1 + i v_1 t + \frac{i^2 v_2}{2!} t^2 + \dots + \frac{i^n v_n}{n!} t^n + O(t^n)$$

2) випливає з 1), оскільки характеристична функція є аналітичною по параметру t , і уявна одиниця – просто фіксований параметр.

Якщо X – дискретна випадкова величина, що задається табличкою

$$X \sim \begin{pmatrix} x_1 & \dots & x_s \\ p_1 & \dots & p_s \end{pmatrix}, \quad \varphi_X(t) = \sum_{j=1}^s e^{itx_j} p_j,$$

$$\frac{d^k \varphi_X(t)}{dt^k} = \sum_{j=1}^s \frac{d^k \left(\sum_{j=1}^s e^{itx_j} p_j \right)}{dt^k} = \sum_{j=1}^s i^k \cdot x_j^k \cdot e^{itx_j} p_j = i^k M(X^k \cdot e^{itX})$$

Додаткові результати :

1. Використаємо наступну теорему математичного аналізу

$$\begin{aligned} \lim_{n \rightarrow \infty} f_n(x) &\rightarrow f(x) \\ \forall n \quad \psi(x) : |f_n(x)| &\leq \psi(x) \\ \int_{-\infty}^{\infty} \psi(x) dx &< \infty \Rightarrow \\ \Rightarrow \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) dx &= \int_{-\infty}^{\infty} \lim_{n \rightarrow \infty} f_n(x) dx. \end{aligned}$$

Цю теорему будемо використовувати у наступному вигляді:

$$|f_n(x, i)| \leq \psi(x)$$

$g(x) \geq 0$ – ще одна невід’ємна функція

$$\begin{aligned} \int_{-\infty}^{\infty} \psi(x) g(x) dx &< \infty \Rightarrow \\ \Rightarrow \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x, i) g(x) dx &= \int_{-\infty}^{\infty} \lim_{n \rightarrow \infty} f_n(x, i) g(x) dx = \int_{-\infty}^{\infty} f(x, i) g(x) dx. \end{aligned}$$

ЛЕКЦІЯ 40

Є випадкова величина. X , задамо індикаторну функцію

$$1_{X \geq 0} = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad 1_{X < 0} = \begin{cases} 0, & x \geq 0 \\ 1, & x < 0 \end{cases}$$

$$1_{X \geq 0} + 1_{X < 0} = 1$$

$$X = (1_{X \geq 0} + 1_{X < 0})X$$

Означення:

$$X^+ = X \cdot 1_{X \geq 0}$$

$$X^- = |X| \cdot 1_{X \leq 0}$$

$$X = X^+ - X^- \Rightarrow MX = MX^+ - MX^-, M|X| = MX^+ + MX^-.$$

MX^+	MX^-	
--------	--------	--

∞	∞	MX невизначено
∞	$< \infty$	$MX = \infty$
$< \infty$	∞	$MX = -\infty$
$< \infty$	$< \infty$	$ MX < \infty, M X < \infty$

Таким чином, якщо $|MX| < \infty \Rightarrow M|X| < \infty$.

Наступний додатковий результат

$$M|X|^n < \infty \Rightarrow M|X|^k < \infty, \quad k \leq n,$$

Виведення для неперервного випадку :

$$M|X|^n = \int_{|x|<1} |x|^n f(x) dx + \int_{|x|\geq 1} |x|^n f(x) dx < \infty$$

$$M|X|^k = \int_{|x|<1} |x|^k f(x) dx + \int_{|x|\geq 1} |x|^k f(x) dx < \infty$$

де $n \geq k$

$M|X|^n < \infty \Rightarrow M|X|^k < \infty$, оскільки $\int_{|x|<1} |x|^n f(x) dx, \int_{|x|<1} |x|^k f(x) dx$ – обмежені, а

з обмеженості $\int_{|x|\geq 1} |x|^n f(x) dx$ випливає обмеженість $\int_{|x|\geq 1} |x|^k f(x) dx$.

Доведення пункту а 4 властивості (неперервний випадок, дискретний випадок $s = \infty$ доводиться аналогічно).

Використовуємо перший додатковий результат:

$$\forall_n |f_n(x, i)| < \psi(x)$$

$$\int_{-\infty}^{\infty} \psi(x) g(x) dx < \infty$$

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x, i) g(x) dx = \int_{-\infty}^{\infty} \lim_{n \rightarrow \infty} (f_n(x, i) g(x)) dx.$$

Використаємо метод математичної індукції

$$\varphi_X^{(0)}(t) = M e^{itX}$$

Нехай формула виконується для k , тож покажемо, що вона виконується для $k + 1$.

$$\varphi_x^{(k+1)} = \lim_{h \rightarrow 0} \frac{\varphi_x^{(k)}(t+h) - \varphi_x^{(k)}(t)}{h} =$$

За означенням похідної:

$$= \frac{i^k M(X^k e^{i(t+h)X}) - i^k M(X^k e^{itX})}{h} = \frac{i^k M(X^k e^{itX} (e^{ihX} - 1))}{h}.$$

Тоді для неперервного випадку отримуємо:

$$\lim_{h \rightarrow 0} i^k \int_{-\infty}^{\infty} x^k e^{itx} \frac{e^{ihx} - 1}{h} f(x) dx$$

$$g(x) = f(x).$$

(Використовуємо формулу:

$$M\varphi(X, i) = \int_{-\infty}^{\infty} \varphi(x, i) f(x) dx.$$

$$f_n(x, i) = x^k e^{itx} \frac{e^{ihx} - 1}{h} =$$

Примітка: $n = \frac{1}{h}$. Впливає $\lim_{n \rightarrow \infty} n = \lim_{h \rightarrow \infty} \frac{1}{h}$.

Знаходимо функцію $\psi(x) \geq \forall_n |f_n(x, i)|$

$$= |x|^k \cdot |e^{itx}| \cdot \left| \frac{e^{ihx} - 1}{h} \right| = \frac{|x|^k}{|h|} |e^{ihx} - 1| = \frac{|x|^k}{|h|} \cdot \left| \int_0^{hx} de^{iu} \right| \leq \frac{|x|^k}{|h|} \int_0^{|h||x|} |ie^{iu}| du = \frac{|x|^k}{|h|}.$$

$$|h||x| = \frac{|x|^{k+1}}{|h|} \cdot |h| = |x|^{k+1}.$$

Покажемо, що інтеграл $\int_{-\infty}^{\infty} \psi(x) g(x) dx < \infty$ (обмежений).

$$\int_{-\infty}^{\infty} |x|^{k+1} f(x) dx = M|X|^{k+1}$$

$$(M\varphi(X) = \int_{-\infty}^{\infty} \varphi(x) f(x) dx)$$

$$|v_n| < \infty$$

$$M|X|^n < \infty$$

$$k+1 \leq n \Rightarrow M|X|^{k+1} < \infty$$

$$\varphi^{(k+1)}(t) = i^k \int_{-\infty}^{\infty} x^k e^{itx} \lim_{h \rightarrow 0} \frac{e^{ihx} - 1}{h} f(x) dx =$$

Для того, щоб взяти ліміт використаємо правило Лопіталя

$$= i^k \int_{-\infty}^{\infty} x^k e^{itx} \left(ix \lim_{h \rightarrow 0} e^{ihx} \right) f(x) dx =$$

$$= i^{k+1} \int_{-\infty}^{\infty} x^{k+1} e^{itx} f(x) dx = i^{k+1} M(X^{k+1} e^{itx}).$$

Приклад 1. Біноміальний розподіл. Як ми пам'ятаємо, його задано ймовірностями

$$p(m) = P(X = m) = C_n^m p^m q^{n-m}, \text{ де } q = 1 - p, m = \overline{0, n}.$$

Характеристична функція дорівнює:

$$\phi_X(t) = M e^{itX} = \sum_{m=0}^n C_n^m e^{itm} p^m q^{n-m} = \sum_{m=0}^n C_n^m (e^{it} p)^m q^{n-m} = (e^{it} p + q)^n.$$

Тут ми використали біном Ньютона.

Приклад 2. Пуассонівський розподіл. У цьому випадку

$$p(m) = P(X = m) = \frac{a^m}{m!} e^{-a}, a > 0, m = \overline{0, \infty}.$$

Обчислюємо характеристичну функцію:

$$\phi_X(t) = \sum_{m=0}^{\infty} e^{itm} \frac{a^m}{m!} e^{-a} = e^{-a} \sum_{m=0}^{\infty} \frac{(ae^{it})^m}{m!} = e^{-a} \cdot e^{ae^{it}} = e^{a(e^{it}-1)}.$$

Приклад 3. Вироджений розподіл:

$$P(X = c) = 1;$$

$$\phi_X(t) = e^{itc}.$$

Приклад 4. Рівномірний розподіл на $[a, b]$. Щільність розподілу має вигляд:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b]; \\ 0, & x \notin [a, b]. \end{cases}$$

Характеристична функція дорівнює:

$$\phi_X(t) = \frac{1}{b-a} \int_a^b e^{itx} dx = \frac{1}{it(b-a)} \int_a^b de^{itx} = \frac{e^{itb} - e^{ita}}{it(b-a)}.$$

Приклад 5. Нехай випадкова величина X має експоненційний (показниковий) розподіл з параметром $\lambda > 0$, тобто щільність розподілу має вигляд:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Характеристична функція експоненційного розподілу дорівнює:

$$\phi_X(t) = \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - it)x} dx = \frac{\lambda}{\lambda - it}.$$

Приклад 6. Нормальний розподіл. Спочатку розглянемо нормований нормальний розподіл, коли щільність розподілу X задано формулою:

$$f(x) = n(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

Характеристична функція дорівнює

$$\varphi_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx.$$

Беремо похідну по t від характеристичної функції

$$\begin{aligned} \varphi_X^{(1)}(t) &= \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{itx - \frac{x^2}{2}} dx = \frac{i}{\sqrt{2\pi}} \left(- \int_{-\infty}^{+\infty} d e^{itx - \frac{x^2}{2}} + it \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx \right) = \\ &= -t \varphi_X(t). \quad \left(\int_{-\infty}^{+\infty} d e^{itx - \frac{x^2}{2}} = 0 \right) \end{aligned}$$

Примітка. $\int_{-\infty}^{+\infty} x e^{itx - \frac{x^2}{2}} dx = - \int_{-\infty}^{+\infty} ((it - x) - it) e^{itx - \frac{x^2}{2}} dx = - \int_{-\infty}^{+\infty} (it - x) e^{itx - \frac{x^2}{2}} dx + \int_{-\infty}^{+\infty} it e^{itx - \frac{x^2}{2}} dx.$

В першому інтегралі робимо заміну:

$$itx - \frac{x^2}{2} = y \Rightarrow dy = (it - x) dx \Rightarrow$$

$$\int_{-\infty}^{+\infty} (it - x) e^{itx - \frac{x^2}{2}} dx = \int_{-\infty}^{+\infty} e^y dy = \int_{-\infty}^{+\infty} d e^y = \int_{-\infty}^{+\infty} d e^{itx - \frac{x^2}{2}}.$$

Отже, дійсно

$$\int_{-\infty}^{+\infty} x e^{itx - \frac{x^2}{2}} dx = - \int_{-\infty}^{+\infty} d e^{itx - \frac{x^2}{2}} + it \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx.$$

$$\varphi_X^{(1)}(t) = -t \varphi_X(t), \quad \varphi_X(0) = 1.$$

$$\varphi_X(t) = e^{-\frac{t^2}{2}}.$$

ЛЕКЦІЯ 41

Знайдемо характеристичну функцію випадкової величини Y з загальним нормальним розподілом $n(y, \nu, \sigma)$.

$$Y = \sigma X + \nu, \quad \sigma \geq 0$$

$$MY = \nu, \quad DY = \sigma^2$$

Характеристична функція випадкової величини Y

$$\varphi_Y(t) = e^{it\nu} \varphi_X(t\sigma) = e^{it\nu} \cdot e^{-\frac{t^2\sigma^2}{2}} = e^{it\nu - \frac{t^2\sigma^2}{2}}.$$

Примітка. Використано властивість 2 характеристичної функції

$$Y = aX + b, \quad \varphi_Y(t) = e^{itb} \varphi_X(at).$$

Покажемо, що випадкова величина Y має нормальний розподіл.

Теорема (без доведення). Для довільної випадкової величини існує взаємно однозначна відповідність між її функцією розподілу і її характеристичною функцією

$$F(x) \leftrightarrow \varphi_X(t).$$

Примітка. Доведемо для частинного випадку.

X – неперервна випадкова величина, а $f(x)$ – неперервна на всій числовій осі.

Маємо:

$$\begin{aligned} \varphi_X(t) &= \int_{-\infty}^{\infty} e^{itx} f(x) dx \\ f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt \end{aligned}$$

– пряме і зворотнє перетворення Фур'є.

В математичному аналізі доводиться, в цьому випадку між $f(x)$ і $\varphi_X(t)$ існує взаємно однозначна відповідність $f(x) \leftrightarrow \varphi_X(t)$.

Якщо в другий інтеграл замість $\varphi_X(t)$ підставити

$$\varphi_X(t) = e^{it\nu - \frac{t^2\sigma^2}{2}},$$

то ми отримаємо:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\nu)^2}{2\sigma^2}}$$

$$m_X(t) = e^{t\nu + \frac{\sigma^2 t^2}{2}}$$

Отже, випадкова величина Y має нормальний розподіл.

***Четверта форма збіжності

Розглянемо нескінченну послідовність випадкових величин $X_1, \dots, X_n, \dots$ та відповідний їм ряд функцій розподілу $F_1(x), \dots, F_n(x), \dots$.

Означення. Послідовність випадкових величин $X_1, \dots, X_n, \dots$ збігається за розподілом до випадкової величини X , якщо ряд їх функцій розподілу $F_1(x), \dots, F_n(x), \dots$ слабо збігається до функції розподілу $F(x)$ ($F(x)$ – функція розподілу випадкової величини X).

Означення. $F_1(x), \dots, F_n(x)$ слабо збігається до $F(x)$, якщо є звичайна збіжність для всіх значень аргументу x , де $F(x)$ – неперервна: $F_n(x) \rightarrow F(x)$ при $n \rightarrow \infty$.

З цього означення випливає, що якщо в точках a і b , функція $F(x)$ неперервна, тоді існує ліміт

$$\exists \lim_{n \rightarrow \infty} P(a \leq X_n < b) = F(b) - F(a).$$

Маємо 2 теореми без доведення.

Теорема 1 (пряма теорема). Маємо послідовність випадкових величин $X_1 \dots X_n$, яка слабо збігається до випадкової величини X , якщо $F_n(x)$ слабо збігається до $F(x)$, тоді їх ряд характеристичних функцій збігається до характеристичної функції $\varphi_X(t)$ випадкової величини X : $\varphi_n(t) \rightarrow \varphi_X(t)$.

Примітка. На якісному рівні справедливості цієї теореми впливає з теореми про взаємно однозначну відповідність між функцією розподілу і характеристичною функцією випадкової величини.

Теорема 2 (обернена теорема). Якщо ряд характеристичних функцій $\varphi_n(t)$, $n = \overline{1, \infty}$ збігається до $\varphi_X(t)$, то ряд функцій розподілу $F_1(x), \dots, F_n(x)$ слабо збігається до функції розподілу $F(x)$ випадкової величини X .

***Центральна гранична теорема, скалярний випадок

Розглянемо послідовність незалежних однаково розподілених випадкових величин $X_1, \dots, X_n, \dots$ кожна з яких $|MX_j| = |\nu| < \infty, DX_j = \sigma^2 < \infty, j = \overline{1, \infty}$.

Примітка. Так як X_j мають однаковий розподіл, то вони мають однакові математичне сподівання та дисперсію.

Розглянемо суму випадкових величин $\sum_{j=1}^n X_j$ і пронормуємо її:

$$Z_n = \frac{\sum_{j=1}^n (x_j - \nu)}{\sigma\sqrt{n}}.$$

Покажемо, що Z_n при $n \rightarrow \infty$ за розподілом збігається до випадкової величини X , що має нормований нормальний розподіл $n(x, 0, 1)$.

Доведення. Нехай $\bar{X}_j = X_j - \nu$, тоді $M\bar{X}_j = 0$,

$$M(\bar{X}_j)^2 = M(X_j - \nu)^2 = DX_j = \sigma^2.$$

Використовуючи пункт 2 властивості 4 характеристичної функції для $n = 2$, отримаємо

$$\varphi_{\bar{X}_j}(t) = 1 + \frac{i\nu_1}{1!}t + \frac{i^2\nu_2}{2!}t^2 + o(t^2)$$

$\nu_1 = 0, \nu_2 = \sigma^2$, отже

$$\varphi_{\bar{X}_j}(t) = 1 - \frac{\sigma^2}{2!}t^2 + o(t^2).$$

Використовуючи властивість 2, в знайдений ряд замість t підставимо $\frac{t}{\sigma\sqrt{n}}$:

$$\varphi_{\frac{\bar{X}_j}{\sigma\sqrt{n}}}(t) = 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{\sigma^2 n}\right).$$

Знайдемо

$$\varphi_{\sum_{j=1}^n \bar{X}_j} \left(\frac{t}{\sigma\sqrt{n}} \right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{\sigma^2 n} \right) \right)^n.$$

Примітка. X_j однаково розподілені та незалежні, тоді всі функції від X_j також незалежні (дивись другий допоміжний результат). Тоді за третьою властивістю характеристична функція суми незалежних випадкових величин дорівнює добутку їх характеристичних функцій.

Спрямуємо $n \rightarrow \infty$, тоді існує наступна границя:

$$\lim_{n \rightarrow \infty} \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{\sigma^2 n} \right) \right)^n = \lim_{n \rightarrow \infty} \left(1 - \frac{t^2}{2n} \right)^n = e^{-\frac{t^2}{2}}$$

(довести самостійно)

Центральна теорема доведена.

Примітка. Використана обернена теорема і теорема про взаємно однозначну відповідність між $F(x)$ та $\varphi_X(x)$.

*****Характеристична функція багатовимірної випадкової величини**

Попередні означення та результати, одержані для одновимірних випадкових величин, природно розповсюджуються на випадкові вектори.

Нехай $X = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix}$ — k -вимірний випадковий вектор-стовпець, а $x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$ —

відповідний вектор дійсних детермінованих змінних.

Нехай $t = \begin{pmatrix} t_1 \\ \vdots \\ t_k \end{pmatrix}$ — вектор-рядок дійсних змінних, кожна з яких приймає значення

від $-\infty$ до $+\infty$.

Означення. Характеристичною функцією випадкового вектора X називається числова скалярна функція дійсного аргументу

$$\varphi_X(t) = M e^{i(t,X)} = M e^{i \sum_{j=1}^k t_j X_j} = \left| (a, b) = \sum_{j=1}^k a_j b_j \right| = M \prod_{j=1}^k e^{it_j X_j}.$$

Наслідок 1. Якщо випадкові величини X_j незалежні, то

$$\varphi_X(t) = \prod_{j=1}^k \varphi_{X_j}(t_j)$$

(властивість 1 характеристичної функції).

Наслідок 2. Характеристична функція обмежена за модулем

$$|Me^{i(t,X)}| \leq M |e^{i(t,X)}| = 1$$

Довести самим, використовуючи формулу Ейлера.

Має місце аналог 4-ої властивості характеристичної функції в скалярному варіанті.

Теорема. Існує взаємно однозначна відповідність між k -вимірними характеристичними функціями k -вимірної випадкової величини та її k -вимірною функцією розподілу : $\varphi_X(t) \leftrightarrow F(x)$.

Примітка. Доведення для випадку, коли існує всюди неперервна k -вимірна функція щільності. Доводиться з використанням прямого і зворотнього перетворення Фур'є для числової скалярної функції f k дійсних аргументів.

Мають також місце пряма і зворотня теорема, яка формулюється ідентично, як і в скалярному випадку, але

$$F_n(x) = F_n(x_1 \dots x_k)$$

$$\varphi_X(t) = \varphi_{X_1, \dots, X_k}(t_1, \dots t_k)$$

Нехай $X_1 \dots X_k$ має k -вимірний нормальний розподіл. Відомо, що її характеристична функція має вигляд

$$\varphi_{X_1, \dots, X_k}(t_1, \dots t_k) = e^{i \sum_{j=1}^k t_j M X_j - \frac{(Bt, t)}{2}},$$

де B – коваріаційна матриця випадкових величин $X_1 \dots X_k$,

$$B = \left(\text{cov}(X_i, X_j) \right)_{i,j=1}^k.$$

Крім того, (Bt, t) – скалярний добуток векторів Bt та t

ЛЕКЦІЯ 42

Розглянемо послідовність $(X(1), X(2) \dots X(n))$ незалежних однаково розподілених k -вимірних випадкових векторів

$$X(l) = \begin{pmatrix} X_1(l) \\ \vdots \\ X_k(l) \end{pmatrix}$$

$$\forall j \quad X(j) = \begin{pmatrix} X_1(j) \\ \vdots \\ X_k(j) \end{pmatrix}$$

Два випадкові вектори $X(j), X(l)$ — **однаково розподілені**, якщо k -вимірна функція розподілу компонент випадкового вектори $X(j)$ дорівнює k -вимірній функції розподілу компонент випадкового вектора $X(l)$.

n випадкових вектори $X(j) \quad j = \overline{1, n}$ **незалежні**, якщо незалежні всі наступні можливі n -вимірні випадкові величини:

$$(X_1(1) \vee X_1(2) \vee \dots \vee X_1(k)), (X_2(1) \vee X_2(2) \vee \dots \vee X_2(k)), \dots, \\ (X_n(1) \vee X_n(2) \vee \dots \vee X_n(k))$$

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} = \begin{pmatrix} MX_1(j) \\ \vdots \\ MX_k(j) \end{pmatrix}.$$

Нехай B -коваріаційна матриця випадкових величин., що складають випадковий вектор $X(l), l = \overline{1, n}$.

$$B = \left(cov \left(X_i(l), X_j(l) \right) \right):$$

Векторна центральна гранична теорема: Нехай $X(1), X(2), \dots, X(n), \dots$ — послідовність k -вимірних незалежних, однаково розподілених випадкових векторів з

вектором математичних сподівань $v = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$ та матрицею коваріацій B , елементи якої

скінченні. Тоді послідовність векторів $Z_n = \frac{\sum_{j=1}^k (X(j) - v)}{\sqrt{n}}$ слабо збігається (збігається за розподілом) до k -вимірного випадкового вектора Z , що має нормальний розподіл

$n(z, \bar{0}^T, B)$, тобто нормальний розподіл з нульовим вектором математичних сподівань та матрицею коваріацій його випадкових аргументів B .

Методологія доведення. Характеристична функція випадкових компонент вектора Z має вигляд $e^{-\frac{(Bt, t)}{2}}$ (загальний вигляд характеристичної функції багатовимірного нормального розподілу наступний:

$$\varphi_{X_1, \dots, X_k}(t_1, \dots, t_k) = e^{i \sum_{j=1}^k t_j M X_j - \frac{(Bt, t)}{2}}$$

За аналогією з одновимірним випадком показується, що існує ліміт:

$$\lim_{n \rightarrow \infty} \varphi_{Z_n(t)} = e^{-\frac{(Bt, t)}{2}}$$

Далі використовується теорема про взаємно однозначну відповідність між характеристичною функцією та k -вимірною функцією розподілу та обернена теорема.

ЛІТЕРАТУРА

1. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. – К.: Вища школа, 1979. – 408с.
2. Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. – М.: Наука, 1985. – 320с.
3. Смирнов Н.В., Дунин – Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1965. – 512с.
4. Худсон Д. Статистика для физиков. – М.: Мир, 1967. – 242с.
5. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648с.
6. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматгиз, 1963. – 500с.
7. Коваленко И.Н., Филиппова А.А. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1982. – 256с.
8. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и её инженерные приложения. – М.: Наука, 1991. – 384с.
9. Свешников А.А. Прикладные методы теории случайных функция. – М.: Наука, 1968. – 463с.
10. Анисимов В.В., Закусило О.К., Донченко В.С. Элементы теории массового обслуживания и асимптотического анализа систем. – К.: Вища школа, 1987. – 248с.
11. Боровиков В.П., Ивченко Г.И. Прогнозирование в системе STATISTICA в среде WINDOWS. – М.: Финансы и статистика, 1999. – 384с.
12. Конышева Л.К., Назаров Д.М. Основы теории нечётких множеств. – СПб.: Питер, 2011. – 192с.

ЗМІСТ

ЛЕКЦІЯ 1.....	3
ЛЕКЦІЯ 2.....	6
ЛЕКЦІЯ 3.....	9
ЛЕКЦІЯ 4.....	12
ЛЕКЦІЯ 5.....	15
ЛЕКЦІЯ 6.....	18
ЛЕКЦІЯ 7.....	21
ЛЕКЦІЯ 8.....	27
ЛЕКЦІЯ 9.....	30
ЛЕКЦІЯ 10.....	35
ЛЕКЦІЯ 11.....	39
ЛЕКЦІЯ 12.....	41
ЛЕКЦІЯ 13.....	44
ЛЕКЦІЯ 14.....	47
ЛЕКЦІЯ 15.....	51
ЛЕКЦІЯ 16.....	54
ЛЕКЦІЯ 17.....	57
ЛЕКЦІЯ 18.....	60
ЛЕКЦІЯ 19.....	63
ЛЕКЦІЯ 20.....	69
ЛЕКЦІЯ 21.....	74
ЛЕКЦІЯ 22.....	75
ЛЕКЦІЯ 23.....	79
ЛЕКЦІЯ 24.....	83
ЛЕКЦІЯ 25.....	88
ЛЕКЦІЯ 26.....	91
ЛЕКЦІЯ 27.....	95
ЛЕКЦІЯ 28.....	99
ЛЕКЦІЯ 29.....	102
ЛЕКЦІЯ 30.....	106
ЛЕКЦІЯ 31.....	108
ЛЕКЦІЯ 32.....	111
ЛЕКЦІЯ 33.....	115

ЛЕКЦІЯ 34.....	120
ЛЕКЦІЯ 35.....	123
ЛЕКЦІЯ 36.....	127
ЛЕКЦІЯ 37.....	129
ЛЕКЦІЯ 38.....	132
ЛЕКЦІЯ 39.....	135
ЛЕКЦІЯ 40.....	140
ЛЕКЦІЯ 41.....	145
ЛЕКЦІЯ 42.....	150
ЛІТЕРАТУРА	152