

Міністерство освіти і науки України
Чернівецький національний університет
імені Юрія Федьковича

Василь Кушнірчук

ТЕОРІЯ ЙМОВІРНОСТЕЙ

Підручник



Чернівці
Чернівецький національний університет
імені Юрія Федьковича
2025

УДК 519.21(075.8)
К 965

Друкується за ухвалою Вченої ради
Чернівецького національного університету імені Юрія Федьковича
(протокол № 15 від 24 листопада 2025 р.)

Рецензенти :

Лукашів Т.О., кандидат фіз-мат. наук, доцент,
науковий співробітник у групі клінічної та
трансляційної інформатики Люксембурзького
центру системної біомедицини
Люксембурзького університету;

Шинкарик М.І, кандидат фіз-мат. наук, доцент,
доцент кафедри прикладної математики
Західноукраїнського національного університету.

Кушнірчук В.Й.

К 965 Теорія ймовірностей : підручник / В.Й. Кушнірчук. Чернівці :
Чернівецьк. нац. ун-т ім. Ю. Федьковича, 2025. 284 с.
ISBN 978-617-8703-38-7

Підручник охоплює основні розділи теорії ймовірностей. Написаний на основі матеріалів курсу „Теорія ймовірностей та математична статистика”, який автор читає студентам факультету математики та інформатики Чернівецького національного університету.

Для студентів вищих навчальних закладів, які навчаються за спеціальностями „Прикладна математика”, „Комп’ютерні науки”, „Системний аналіз та наука про дані” та ін., а також для тих, хто бажає самостійно вивчати теорію ймовірностей.

УДК 519.21(075.8)

© Чернівецький національний університет
імені Юрія Федьковича, 2025
© Кушнірчук В.Й., 2025

ISBN 978-617-8703-38-7

ВСТУП

Теорія ймовірності – це наука, предметом вивчення якої є встановлення закономірностей у масових випадкових явищах (подіях). Вона слугує основою для статистики, машинного навчання, фінансового аналізу та багатьох інших галузей, що робить її однією з ключових дисциплін сучасної науки.

Витоки теорії ймовірностей сягають XVI–XVII століть, коли європейські математики почали систематично досліджувати випадкові процеси, зокрема в контексті азартних ігор. Одними з перших, хто заклав основи дисципліни, були італійський математик Джироламо Кардано та французькі вчені Блез Паскаль і П'єр Ферма. Їхня переписка в 1654 році, присвячена проблемі розподілу ставок у грі, вважається початком формального вивчення ймовірностей. У цей же період Християн Гюйгенс опублікував працю „Про розрахунки в азартних іграх” (1657), яка стала першим систематичним викладом основ теорії.

У XVIII столітті теорія ймовірностей почала набувати сучасних рис завдяки працям Якоба Бернуллі, який сформулював закон великих чисел, та Абрахама де Муавра, який розробив основи теорії розподілів і ввів поняття нормального розподілу. У XIX столітті значний внесок зробили П'єр-Сімон Лаплас, який узагальнив багато ідей у своїй праці „Аналітична теорія ймовірностей” (1812), та Карл Фрідріх Гаус, чиї наукові роботи з нормального розподілу стали основою для статистичних методів.

У XX столітті теорія ймовірностей отримала строге математичне підґрунтя завдяки аксіоматичному підходу, розробленому Андрієм Колмогоровим. Його книга „Основи теорії ймовірностей” (1933) стала переломною, запровадивши аксіоми, які формалізували дисципліну та зробили її частиною сучасної математики. Інші видатні вчені, такі як Еміль Борель, Норберт Вінер і Джон фон Нейман, розвинули

теорію в контексті стохастичних процесів, теорії інформації та квантової механіки.

Українські математики також зробили вагомий внесок у розвиток теорії ймовірностей. Одним із найвидатніших є Віктор Глушков, який, хоча й більш відомий як піонер кібернетики, активно застосовував імовірнісні методи в теорії автоматів і комп'ютерних систем. Значний внесок у теорію випадкових процесів зробив Йосип Бернштейн, який працював над стохастичними диференціальними рівняннями. У другій половині XX століття український математик Анатолій Скороход розробив фундаментальні результати в галузі теорії випадкових процесів, зокрема в теорії марковських процесів і стохастичного аналізу. Його праці стали класикою та використовуються в усьому світі. Вагомі результати в теорії ймовірностей одержали Борис Гнеденко, Михайло Ядренко, Володимир Королюк, Юрій Козаченко. Їхні наукові роботи знаходять застосування в багатьох актуальних галузях.

Теорія ймовірностей є не лише абстрактною математичною дисципліною, але й потужним інструментом для аналізу реального світу. Вона допомагає прогнозувати результати експериментів, оцінювати ризики, оптимізувати процеси та приймати обґрунтовані рішення в умовах невизначеності. Цей підручник покликаний познайомити читачів з основними поняттями, методами та застосуваннями теорії ймовірностей, а також надихнути на подальше її дослідження.

ВІД АВТОРА

Автор вважає приємним обов'язком висловити щиру вдячність професору Ігорю Малику за увагу до цієї книги і корисні поради при її написанні, а також глибоку повагу своїм викладачам теорії ймовірностей Антону Садов'яку та Марку Букатарю.

РОЗДІЛ 1

ВИПАДКОВІ ПОДІЇ

§1. Стохастичний експеримент. Випадкова подія. Частота випадкової події

Уявімо собі такий експеримент: нехай багато разів підкидається кубик в однакових умовах. У кожному підкиданні кількість очок, що випадають на верхній грані кубика – це результат експерименту. Цим результатом можуть бути числа 1, 2, 3, 4, 5, 6.

Якщо такий експеримент проводити багато разів в одних і тих самих умовах, то такий експеримент називається **стохастичним**. Результат кожного окремого підкидання передбачити неможливо, але множину результатів можна описати. Будемо описувати елементарні результати експерименту, які вже не розбиваються (не інтерпретуються) дрібнішими (елементарнішими). З кожним елементарним результатом пов'яжемо події – „випала одиниця”, „випала двійка”, ..., „випала шістка”.

Множина цих елементарних подій називається **простором елементарних подій** і позначається Ω .

На простір елементарних подій накладається умова – у результаті експерименту може з'явитися тільки один елемент простору Ω . З елементарних подій утворюються більш складні події. Наприклад, „на верхній грані парне число”, „на верхній грані число більше від 3” і ін. Кожна з елементарних подій може відбутися або ні, але одна з них точно буде мати місце (відбудеться). Такі події будемо називати **випадковими подіями**.

При підкиданні монети, наприклад, елементарні події „випав герб” і „випала цифра”. Позначимо їх, відповідно, $\Gamma = \{\text{„випав герб”}\}$ та $\Pi = \{\text{„випала цифра”}\}$, тоді простір $\Omega = \{\Gamma, \Pi\}$.

Не слід думати, що простір елементарних подій обов’язково скінченний. Розглянемо такий експеримент: підкидається монета, поки на верхній грані не випаде герб. Простір елементарних подій у цьому експерименті $\Omega = \{\Gamma, \Pi\Gamma, \Pi\Pi\Gamma, \Pi\Pi\Pi\Gamma, \dots\}$. Тут зліченна кількість елементарних подій.

А при вимірюванні довжини простором елементарних подій є відрізок, якщо вважати, що будь-яке дійсне число може бути результатом вимірювання. Тут континуальна кількість елементарних подій.

Якщо експеримент проводити масово, нехай n разів, і відмічати кількість експериментів m , у яких деяка зафіксована подія відбувається, то відношення $\frac{m}{n}$ називається **частотою події**. Якщо цю подію позначити A , а $m(A)$ – кількість експериментів, у яких A відбулася, то частота події A у стохастичному експерименті обчислюється за формулою $\nu_n(A) = \frac{m(A)}{n}$. Виявляється, що частота події має властивість стійкості, яка полягає у тому, що існує $\lim_{n \rightarrow \infty} \nu_n(A)$, позначимо її $p(A)$. Це означає, наприклад при підкиданні монети, що частота появи Π при дуже великих n мало відрізняється від $\frac{1}{2}$. Через стійкість частоти події виникла думка, чи не можна частоту використати як характеристику події. Можна, але для цього треба проводити дуже багато складних і часом дуже дорогих, руйнівних і т.д. експериментів. Тому доводиться шукати інші підходи для характеристики випадкових подій.

Нехай задано простір елементарних подій Ω .

Означення. *Випадковою подією у стохастичному експерименті називається будь-яка підмножина простору елементарних подій.*

Елементарні події, які є елементами цієї підмножини, називаються **елементарними подіями**, що спричиняють випадкову подію.

Надалі випадкові події позначаються $A, B, C, \dots$ і т. д. – це підмножини множини Ω .

Сама подія Ω (вона є підмножиною Ω) є **достовірною** (вірогідною) **подією**. Вона так називається тому, що відбувається завжди, в кожному експерименті.

Підмножиною Ω є і порожня множина $\emptyset$. Таку подію будемо називати **неможливою подією**, її не спричиняє жодна з елементарних подій.

Приклад. Нехай розглядається стохастичний експеримент із підкиданням кубика. Тут $\Omega = \{1, 2, 3, 4, 5, 6\}$. Розглянемо подію A – на верхній грані кубика випало число, яке ділиться на 3 без остачі. Тоді $A = \{3, 6\}$. Елементарні події $\{3\}, \{6\}$ спричиняють подію A . Розглянемо подію B – випало парне число. $B = \{2, 4, 6\}$ – підмножина Ω .

У подіях A та B є спільні елементи (елемент $\{6\}$). Це означає, що A та B можуть відбуватися одночасно.

Означення. Події, які можуть відбуватися одночасно, називаються **сумісними подіями**.

Опишемо ще подію C – на верхній грані випало число, яке не більше двох: $C = \{1, 2\}$. Події A і C не можуть відбуватися одночасно (не мають однакових елементарних подій).

Означення. Події, які не можуть відбуватися одночасно, називаються **несумісними подіями**.

У експерименті з підкиданням монети до першої появи герба простір елементарних подій $\Omega = \{\Gamma, ЦГ, ЦЦГ, ЦЦЦГ, ЦЦЦЦГ, \dots\}$. Розглянемо подію A – герб появився після парного числа підкидань. Тоді $A = \{ЦГ, ЦЦЦГ, ЦЦЦЦЦГ, \dots\}$.

Повернемося до розгляду поняття частоти. Події можуть бути сумісні і несумісні. Якщо події A і B несумісні, а це означає, що

$A \cap B = \emptyset$, то легко зрозуміти, що в стохастичному експерименті сума частот події A та події B дорівнює частоті події (A або B):

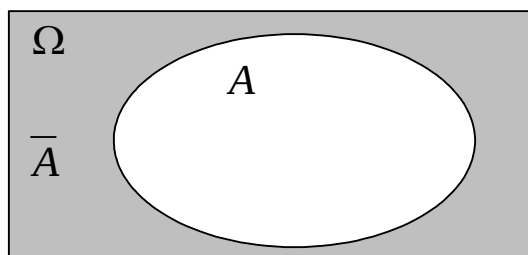
$$\nu_n(A) + \nu_n(B) = \nu_n(A \text{ або } B).$$

Така властивість повинна зберігатися у тієї ознаки, якою будемо характеризувати подію.

Нехай задана подія A .

Означення. *Протилежною до події A називається подія $\bar{A}$, яка відбувається тоді і тільки тоді, коли подія A не відбувається.*

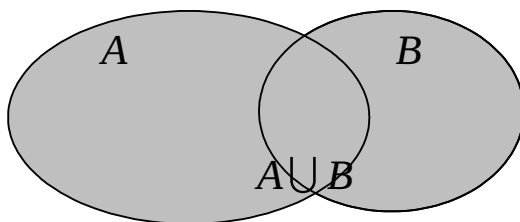
Легко зрозуміти, що коли події A відповідає якась множина, то події $\bar{A}$ відповідає доповнення цієї множини до множини Ω . Це зручно ілюструвати при допомозі діаграм Венна*:



Нехай задано події A і B .

Означення. *Сумою подій A і B називається така подія C , яка відбувається тоді і тільки тоді, коли відбувається хоча б одна з подій A або B .*

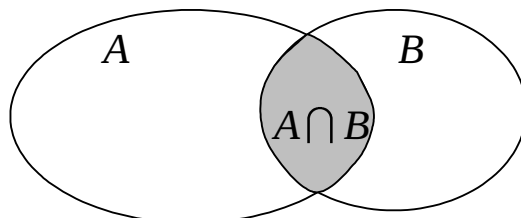
Подія A спричиняється набором подій (множина A), події B відповідає інша множина. Події C відповідає об'єднання множин A і B . Тому природно позначити суму подій (використати для позначення події C) так: $C = A \cup B$.



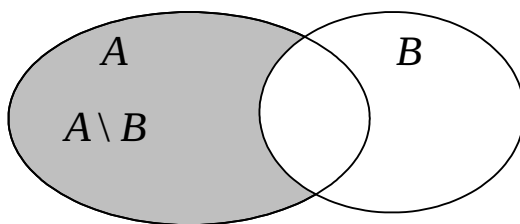
* Джон Венн (John Venn, 1834-1923) – англійський логік і філософ.

Означення. Добутком подій A і B називається подія C , яка відбувається тоді і тільки тоді, коли відбуваються і подія A , і подія B .

Повинні відбутися елементарні події, які спричиняють появу і події A і, події B , тобто перетин подій, тому природно записувати $C = A \cap B$.



Означення. Різницею подій A і B називається подія C , яка полягає в тому, що подія A відбувається і подія B не відбувається. Прийнято записувати $C = A \setminus B = A \cap \bar{B}$.



Тепер можна записати такі очевидні співвідношення:

$\bar{\Omega} = \emptyset$, $\bar{\emptyset} = \Omega$, $A \cup A = A$, $A \cap A = A$, $A \cup \Omega = \Omega$, $A \cap \Omega = A$ і інші.

Оскільки, A – множина елементарних подій, то всі закони алгебри множин мають місце для подій, зокрема:

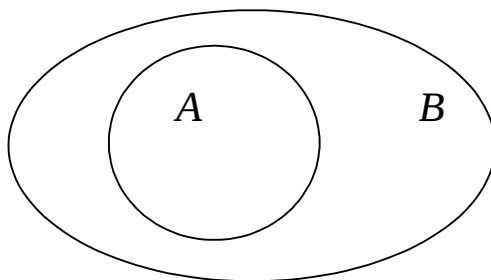
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}, \quad \overline{A \cup B} = \bar{A} \cap \bar{B}.$$

Ці рівності можна довести. Доведемо одну з них, наприклад $\overline{A \cup B} = \bar{A} \cap \bar{B}$.

◀ Припустимо, що відбулася подія $\overline{A \cup B}$. Це означає, що $A \cup B$ не відбулася, а отже, не відбулася ні подія A , ні подія B , тобто відбулася подія $\bar{A} \cap \bar{B}$. Тепер навпаки: нехай відбулася подія $\bar{A} \cap \bar{B}$. Це означає, що не відбулася ні подія A , ні подія B , тобто не відбулася подія $A \cup B$, а тому відбулася подія $\overline{A \cup B}$. ▶

Означення. Кажуть, що подія A викликає (спричиняє) подію B , якщо виконання події A приводить до виконання події B .



Записують це так: $A \subset B$. Якщо подія A спричиняє подію B ($A \subset B$) і подія B спричиняє подію A ($B \subset A$), то події A і B вважають однаковими (рівними), тобто $A = B$.

Очевидно, що A викликає Ω ($A \subset \Omega$), а також $A \subset A \cup B$, $A \cap B \subset B$, $A \cap B \subset A$.

Отже, якщо події A і B несумісні (тобто $A \cap B = \emptyset$), то частота суми подій дорівнює сумі частот: $\nu_n(A \cup B) = \nu_n(A) + \nu_n(B)$. Крім цього, частота $\nu_n(A) \geq 0$, а частота достовірної події $\nu_n(\Omega) = 1$ і частота неможливої події $\nu_n(\emptyset) = 0$. З останніх трьох відношень можна зробити висновок, що частота довільної події $0 \leq \nu_n(A) \leq 1$. Та ознака, якою будемо характеризувати подію, повинна задовольняти ці самі властивості.

Нехай задана послідовність подій $A_1, A_2, \dots, A_n, \dots$

Означення. Верхньою границею послідовності подій (при $n \rightarrow \infty$) називається подія, яка полягає в тому, що відбувається нескінченне число подій послідовності. Ця подія позначається $A^* = \overline{\lim_{n \rightarrow \infty} A_n}$.

Якщо розглядається нескінченна послідовність подій $A_1, A_2, \dots, A_n, \dots$, то завжди існує номер N , такий, що подія з номером більшим за N (хоч одна) завжди відбувається. Тому при кожному m відбувається хоч одна подія $\bigcup_{n=m}^{\infty} A_n$. Отже, $A^* = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$.

Означення. Нижньою границею послідовності подій називається подія, яка полягає в тому, що не відбувається не більше ніж скінченна кількість подій послідовності.

Або інакше: відбуваються всі події послідовності, за винятком, можливо, скінченної кількості. Тому при якомусь m відбувається подія $\bigcap_{n=m}^{\infty} A_n$, а отже, $A_* = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n$. Зрозуміло, що A_* викликає подію A^* , тобто $A_* \subset A^*$. Позначається ця подія $A_* = \varliminf_{n \rightarrow \infty} A_n$.

Якщо події A_* і A^* співпадають (тобто $A_* = A^* = A$ – позначення), то кажуть, що послідовність подій має границею подію A .

У частковому випадку задана послідовність подій може бути монотонною. Припустимо, що $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$, тоді зрозуміло, що $A_* = A^* = A = \bigcup_{n=1}^{\infty} A_n$, бо $\bigcap_{n=m}^{\infty} A_n = A_m$, і можна довести, що $\bigcup_{n=m}^{\infty} A_n = A_m$.

Аналогічно, якщо $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$, то $A_* = A^* = A = \bigcap_{n=1}^{\infty} A_n$. Це також можна строго довести.

Наведемо паралелі між подіями і теоретико-множинною теорією:

		Теорія ймовірностей	Теорія множин
1	Ω	простір елементарних подій, достовірна подія	універсальна множина
2	ω	елемент простору елементарних подій, елементарний результат експерименту	точка множини $\omega \in \Omega$
3	A	подія	підмножина – $A \subset \Omega$
4	$\emptyset$	неможлива подія	порожня множина
5	$\bar{A}$	протилежна подія	доповнення множини A до Ω : $\bar{A} = C_{\Omega}(A)$
6	$A \cup B$	сума подій	об'єднання множин
7	$A \cap B$	добуток подій	перетин множин
8	$A \setminus B$	різниця подій	різниця множин

§2. Класичне означення ймовірності

Нехай задано простір елементарних подій Ω . Припустимо, що він складається не більше ніж зі зліченної кількості елементарних подій (тобто ці елементарні події можна перенумерувати і їх може бути скінченна кількість): $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$. Нехай з якихось міркувань кожному ω_i поставлено у відповідність дійсне число p_i так, що

$$1) \quad p_i \geq 0,$$

$$2) \quad \sum_{i=1}^{\infty} p_i = 1.$$

Означення. Нехай A деяка подія, тобто $A \subset \Omega$. **Ймовірністю події A** називають число $P(A)$, яке, за означенням, дорівнює

$$P(A) = \sum_{i: \omega_i \in A} p_i.$$

Це означення можна назвати **класичним означенням імовірності**. Тут не говорять про те, як ω_i ставиться у відповідність число p_i . Якщо по-різному вибирати ці числа, то можна отримати різні результати.

Приклад 1. Будемо розглядати підкидання правильної монети. При такому однократному експерименті на верхній грані монети може випасти цифра або герб. У цьому випадку $\Omega = \{\text{Ц}, \text{Г}\}$ (тобто $\omega_1 = \text{Ц}$, $\omega_2 = \text{Г}$) – скінченна множина. Кожному з елементів ω_1 і ω_2 треба поставити у відповідність числа з властивостями 1) і 2). Якщо сторони монети рівноправні, то природно вибрати ці числа такими:

$$p_1 = \frac{1}{2}, \quad p_2 = \frac{1}{2}.$$

У цьому експерименті можна розглядати події $\emptyset$, $\{\text{Г}\}$, $\{\text{Ц}\}$, $\{\text{Ц}, \text{Г}\} = \Omega$. Очевидно, що $P(\emptyset) = 0$, бо в $\emptyset$ не входять ω_1 та ω_2 і

$$\sum_{i: \omega_i \in \emptyset} p_i = 0. \quad P(\{\Gamma\}) = \frac{1}{2}, \quad P(\{\Pi\}) = \frac{1}{2}, \quad P(\Omega) = 1.$$

Приклад 2. Розглянемо експеримент, який полягає у підкиданні правильного грального кубика. Простір елементарних подій складається з шести елементарних подій: 1 – „на верхній грані кубика випала цифра 1”, і т.д. Тому $\Omega = \{1; 2; 3; 4; 5; 6\}$. Оскільки всі грані рівноправні, кубик правильний, то приписуємо кожній елементарній події однакові числа з властивостями 1) і 2). Тому $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$.

У цьому експерименті можна розглядати різні події. Всього подій може бути 2^6 . Розглянемо таку подію: A – „на верхній грані з'явилося число, яке при діленні на 3 дає остачу 1”. Тоді $A = \{1; 4\}$, а ймовірність цієї події, згідно з означенням,

$$P(A) = p_1 + p_4 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

Приклад 3. Розглянемо „фальсифікований” кубик. Для того, щоб частіше випадало 6, центр ваги треба змістити вниз, щоб 1 була внизу, а 6 вгорі. Припустимо, що центр ваги зміщено так, що числа, які ставимо у відповідність елементарним подіям, визначено як $\omega_6 \rightarrow p_6 = \frac{7}{36}$, $\omega_1 \rightarrow p_1 = \frac{5}{36}$, а для решти подій $p_2 = p_3 = p_4 = p_5 = \frac{1}{6}$. У цьому експерименті для події A , розглянутої в прикладі 2, імовірність $P(A)$ становить $\frac{11}{36}$.

Як вибирати ці числа p_i , розглядати не будемо, можна багато разів проводити експеримент, проаналізувати частоти і в певний спосіб вибрати потім числа p_i .

Приклад 4. Проводиться підкидання монети доти, поки вона не випаде гербом наверх. Тут

$$\Omega = \{\Gamma, \text{ЦГ}, \text{ЦЦГ}, \text{ЦЦЦГ}, \dots, \text{ЦЦЦ...ЦГ}, \dots\}.$$

Маємо справу з Ω , що містить зліченну кількість точок (елементарних наслідків ω_i). Елементарним наслідкам ставимо у

$$\text{відповідність числа } \omega_1 \rightarrow p_1 = \frac{1}{2}, \quad \omega_2 \rightarrow p_2 = \left(\frac{1}{2}\right)^2, \quad \omega_3 \rightarrow p_3 = \left(\frac{1}{2}\right)^3,$$

$$\omega_4 \rightarrow p_4 = \left(\frac{1}{2}\right)^4, \dots, \quad \omega_{n+1} \rightarrow p_{n+1} = \left(\frac{1}{2}\right)^{n+1}, \dots \text{Перевіримо умови 1) і 2).}$$

Умова 1) виконується, бо всі $p_i \geq 0$. Умова 2) також виконується:

$$\sum_i p_i = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \dots = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1.$$

Розглянемо подію A , яка полягає у тому, що кількість підкидань ділиться на три без остачі. Тоді вона має вигляд $A = \{\text{ЦЦГ}, \text{ЦЦЦЦГ}, \text{ЦЦЦЦЦЦЦГ}, \dots\}$. За означенням, імовірність цієї події

$$P(A) = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^9 + \dots = \frac{\left(\frac{1}{2}\right)^3}{1 - \left(\frac{1}{2}\right)^3} = \frac{1 \cdot 8}{8 \cdot 7} = \frac{1}{7}.$$

Взагалі, якщо задано простір елементарних подій Ω і введена в якийсь спосіб імовірність P (тобто числа $p_1, p_2, \dots$), то кажуть, що задано ймовірнісну модель (Ω, P) .

Частковим випадком сказаного раніше є випадок скінченного Ω (скінченної кількості елементів у цій множині) і рівноможливих елементарних результатів експерименту. У цьому випадку всі p_i однакові, тобто $p_i = p$. Тому ймовірність будь-якої події A така:

$P(A) = \sum_{i:\omega_i \in A} p_i = \sum_{i:\omega_i \in A} p = p \cdot m$, де m – кількість елементарних подій,

які спричиняють подію A (які викликають подію A), або кількість сприятливих подій A елементарних подій. Але сума по всіх i

$\sum_{i:\omega_i \in \Omega} p_i = p \sum_{i:\omega_i \in \Omega} 1 = p \cdot n = 1$, де n – кількість всіх елементарних подій.

Звідси маємо $p = \frac{1}{n}$. Підставляємо в $P(A)$ і одержуємо

$P(A) = p \cdot m = \frac{m}{n}$, тобто

$$P(A) = \frac{m}{n}$$

– **класичне означення ймовірності**: ймовірність події A дорівнює відношенню кількості сприятливих подій A елементарних подій до загальної кількості елементарних подій.

Приклад 5. Із ретельно перемішаної колоди карт (24 листи) навмання виймають 3 карти. Яка ймовірність того, що вийнята одна дама і два тузи?

◀ Ретельно перемішана – означає рівну можливість витягнути довільну карту. Навмання – означає рівну можливість витягнути довільну трійку карт.

У цьому прикладі загальна кількість можливих випадків – це кількість способів, якими можна вибрати 3 карти з 24-х:

$$n = C_{24}^3 = \frac{24!}{3! \cdot 21!} = \frac{22 \cdot 23 \cdot 24}{1 \cdot 2 \cdot 3} = 4 \cdot 23 \cdot 22.$$

Кількість способів вибрати одну даму – 4, а кількість способів вибрати з чотирьох тузів два тузи – C_4^2 . Тому кількість сприятливих для настання події A випадків $m = 4 \cdot C_4^2 = 4 \cdot \frac{4!}{2! \cdot 2!} = 4 \cdot \frac{3 \cdot 4}{2} = 24$.

Тоді $P(A) = \frac{m}{n} = \frac{24}{4 \cdot 22 \cdot 23} = \frac{3}{11 \cdot 23} = \frac{3}{253}$. ▶

§3. Геометрична ймовірність

Якщо вести стрільбу по мішені, то виникає питання про відстань від точки попадання до цілі. Ця відстань може бути довільним числом з деякого відрізка. Тоді простір елементарних подій є континуальним і знаходити імовірності подій за класичним означенням неможливо.

Часто простором елементарних подій є деяка частина або площини, або прямої, або 3-вимірного простору, а події, які розглядаються, мають таку властивість, що їм можна поставити у відповідність деяку „міру”.

Тоді **ймовірністю події** A називають міру A ($A \subset \Omega$), поділену на міру Ω , тобто

$$P(A) = \frac{\text{mes}(A)}{\text{mes}(\Omega)}.$$

Так визначається **геометрична ймовірність**.

Приклад (задача про зустріч).

Двоє домовилися про зустріч в якийсь фіксований час. Домовленість така: зустріч може відбутися протягом часу T (наприклад, однієї години). Перший, хто прийде, чекає другого протягом часу $\tau < T$. Яка ймовірність того, що зустріч відбудеться, якщо моменти приходу учасників зустрічі рівноправні (однаково можливий прихід учасників протягом часу T).

◀ Позначимо: x – момент приходу першого учасника зустрічі (від початку обумовленого часу), y – момент приходу другого учасника зустрічі. Зрозуміло, що $0 \leq x \leq T$, $0 \leq y \leq T$. Тоді елементарним результатом цього експерименту буде пара чисел (x, y) . Який простір елементарних подій? Простором елементарних подій є такі пари чисел (x, y) , що $0 \leq x \leq T$ і $0 \leq y \leq T$, тобто

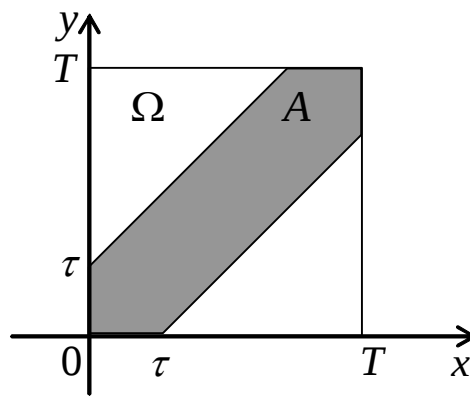
$$\Omega = \{(x, y) : 0 \leq x \leq T, 0 \leq y \leq T\}.$$

Геометрично (на площині) Ω – точки відповідного квадрату. Кожна точка визначає моменти приходу учасників зустрічі. Опишемо подію A – зустріч відбудеться:

$$\Omega \supset A = \{(x, y) : 0 \leq x \leq T, 0 \leq y \leq T, |x - y| \leq \tau\}.$$

Зобразимо множину A , тобто множину точок квадрата Ω , що задовольняють нерівність $|x - y| \leq \tau$, яка рівносильна таким нерівностям:

$$-\tau \leq x - y \leq \tau \Rightarrow \begin{cases} x - y \leq \tau, \\ x - y \geq -\tau, \end{cases} \Rightarrow \begin{cases} x - y \leq \tau, \\ y - x \leq \tau. \end{cases}$$



Як міру на площині будемо використовувати площу. Тому

$$\text{mes } \Omega = T^2 \quad \text{mes } A = T^2 - (T - \tau)^2.$$

Тоді, за означенням, імовірність того, що зустріч відбудеться, дорівнює

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

Якщо б задача формулювалась так: чекають менше ніж τ , то імовірність $P(A)$ була б така сама. ►

Із даних вище означень випливає, що ймовірність будь-якої події невід'ємна $P(A) \geq 0$, не більша одиниці $P(A) \leq 1$, а ймовірність протилежної події $P(\bar{A}) = 1 - P(A)$. Якщо події A і B несумісні (тобто відповідні множини неперетинні $A \cap B = \emptyset$, диз'юнктивні), то $P(A \cup B) = P(A) + P(B)$. Такі самі властивості були й у частоти події. Отже, всі властивості частоти зберігаються для введених подій (простих подій, які вже означені).

§4. Класи подій

Нехай Ω – простір елементарних подій і події, як і раніше, трактуються як підмножини множини Ω . Нехай маємо якісь події $A, B, C, \dots$. Можна розглядати протилежні події $\bar{A}, \bar{B}, \bar{C}, \dots$, мають зміст ще й інші події, наприклад $A \cup C, A \cap B, A \setminus B = A \cap \bar{B}$. Тобто разом з подіями $A, B, C, \dots$ треба розглядати багато інших подій, утворених (сконструйованих) із подій $A, B, C, \dots$ різними операціями.

Означення 1. Клас подій $\mathcal{A}$ називається **алгеброю подій**, якщо виконуються умови:

1) достовірна подія є елементом цього класу $\Omega \in \mathcal{A}$;

2) якщо деяка подія A є елементом класу, то протилежна подія $\bar{A}$ теж є елементом цього класу:

$$A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A};$$

3) якщо події A і B є елементами класу, то їх сума теж є елементом цього класу:

$$A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}.$$

З умови 3, за індукцією, можна встановити, що з $A_1, A_2, \dots, A_n \in \mathcal{A}$ випливає $\bigcup_{i=1}^n A_i \in \mathcal{A}$.

Умову 3 можна замінити на таку еквівалентну умову:

3) якщо події A і B є елементами класу, то їх добуток теж є елементом цього класу:

$$A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}.$$

Легко переконатися, що це так, застосувавши закони де Моргана*:

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \quad \overline{A \cap B} = \bar{A} \cup \bar{B}.$$

* Аугустус де Морган (Augustus De Morgan, 1806-1871) – британський математик та логік.

◀ Доведення того, що з умови 3 випливає умова 3'. Нехай має місце умова 3 і $A, B \in \mathcal{A}$. Тоді, за умовою 2, $\bar{A}, \bar{B} \in \mathcal{A}$, тому, за умовою 3, $\overline{\bar{A} \cup \bar{B}} \in \mathcal{A}$, а тепер, за умовою 2, $\overline{\overline{\bar{A} \cup \bar{B}}} \in \mathcal{A}$. За законом де Моргана, $\overline{\overline{\bar{A} \cup \bar{B}}} = A \cap B \in \mathcal{A}$. А тому $A \cap B \in \mathcal{A}$, тобто має місце умова 3'. ▶

Приклад 1. Найпростіший приклад алгебри: нехай Ω – простір елементарних подій, $\mathcal{A} = \{\Omega, \emptyset\}$. Умови 1, 2, 3 виконуються.

Приклад 2. $\mathcal{A} = \{\Omega, A, \bar{A}, \emptyset\}$, де A – довільна підмножина Ω . Умови 1, 2, 3 виконуються.

Приклад 3. Не є алгеброю $\mathcal{A} = \{\Omega, A, \emptyset\}$.

Означення 2. *Спадковим класом* множини Ω називається клас усіх можливих підмножин множини Ω .

Очевидно, що спадковий клас є алгеброю.

Означення 3. *Клас подій $\mathcal{F}$ називається σ -алгеброю, якщо:*

- 1) достовірна подія належить цьому класу $\Omega \in \mathcal{F}$;
- 2) якщо $A \in \mathcal{F}$, то звідси випливає, що $\bar{A} \in \mathcal{F}$;
- 3) якщо $A_1 \in \mathcal{F}, A_2 \in \mathcal{F}, \dots, A_n \in \mathcal{F}, \dots$, то звідси випливає, що

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}.$$

Відмінність між класами подій $\mathcal{A}$ і $\mathcal{F}$ в умові 3: $\mathcal{F}$ – більш широке (тут сума зліченної кількості подій). Зрозуміло, що всяка σ -алгебра є алгеброю, але не всяка алгебра є σ -алгеброю.

Означення 4. *Клас подій (множин) $\mathcal{M}$ називається монотонним класом, якщо:*

- 1) з умов $A_1 \in \mathcal{M}, A_2 \in \mathcal{M}, \dots, A_n \in \mathcal{M}, \dots$, і $A_1 \subset \dots \subset A_n \dots$ випливає,

$$\text{що } \lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n \in \mathcal{M};$$

- 2) якщо $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \dots$, то $\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n \in \mathcal{M}$.

Отже, монотонний клас замкнений відносно границь монотонних послідовностей.

Приклад 4. Всяка σ -алгебра є монотонним класом, тобто вона замкнена відносно будь-якої послідовності.

Означення 5. Нехай $\mathcal{K}$ – деякий клас подій (множин). Тоді **мінімальною σ -алгеброю** (позначається $\sigma(\mathcal{K})$), яка містить клас $\mathcal{K}$, називається така σ -алгебра, що задовольняє умови

- 1) $\mathcal{K} \subset \sigma(\mathcal{K})$;
- 2) якщо $\mathcal{F}$ – довільна σ -алгебра, яка містить $\mathcal{K}$ ($\mathcal{F} \supset \mathcal{K}$), то $\sigma(\mathcal{K})$ міститься у $\mathcal{F}$, тобто $\sigma(\mathcal{K}) \subset \mathcal{F}$.

Зрозуміло, що мінімальна σ -алгебра існує і це – переріз усіх σ -алгебр, які містять клас $\mathcal{K}$.

Означення 6. **Мінімальним монотонним класом**, який містить клас $\mathcal{K}$ (позначається $\mathcal{M}(\mathcal{K})$), називається такий монотонний клас, що

- 1) $\mathcal{K} \subset \mathcal{M}(\mathcal{K})$;
- 2) якщо $\mathcal{M}$ такий монотонний клас, що містить $\mathcal{K}$, то $\mathcal{M}(\mathcal{K})$ міститься в $\mathcal{M}$: $\mathcal{M} \supset \mathcal{K} \Rightarrow \mathcal{M}(\mathcal{K}) \subset \mathcal{M}$.

Можна довести, що коли $\mathcal{K}$ є алгеброю, то $\sigma(\mathcal{K}) = \mathcal{M}(\mathcal{K})$.

Сформулюємо твердження про зв'язок алгебри і σ -алгебри.

Твердження. Алгебра $\mathcal{A}$ є σ -алгеброю тоді і тільки тоді, коли $\mathcal{A}$ є монотонним класом.

◀ **Необхідність** очевидна, оскільки, якщо алгебра $\mathcal{A}$ є σ -алгеброю, то $\mathcal{A}$ є монотонним класом.

Достатність. Нехай алгебра $\mathcal{A}$ є монотонним класом. Доведемо, що $\mathcal{A}$ є σ -алгеброю. Перевіримо, чи виконуються умови 1–3 з означення σ -алгебри.

1. $\Omega \in \mathcal{A}$.
2. $A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$.

3. За умовою 2, означення монотонного класу з того, що $A_1, A_2, \dots, A_n, \dots \in \mathcal{A}$, випливає $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Утворимо множини $B_1 = A_1, B_2 = A_1 \cup A_2, \dots, B_m = \bigcup_{i=1}^m A_i, \dots$. При кожному m це є елементи алгебри, крім цього $B_1 \subset B_2 \subset B_3 \subset \dots$, тобто це монотонний клас, а тому $\lim_{m \rightarrow \infty} B_m \in \mathcal{A}$, тобто $\lim_{m \rightarrow \infty} B_m = \bigcup_{m=1}^{\infty} B_m = \bigcup_{i=1}^{\infty} A_i$.

Отже, одержали, що $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$. Умову 3 доведено. ►

Зі сформульованого і доведеного твердження випливає такий висновок: з алгебри можна утворити σ -алгебру, якщо до алгебри додати границі всіх монотонних послідовностей.

Наприклад, якщо маємо дійсне число $\{a\}$, то розглянемо інтервали

$$A_n = \left[a, a + \frac{1}{n} \right).$$

Зрозуміло, що $A_n \supset A_{n+1}$ і тоді $\bigcap_{n=1}^{\infty} A_n = \{a\}$.

Елементи всіх класів є подіями.

Означення 7. 1) Нехай задана деяка алгебра $\mathcal{A}$, кожному елементу якої поставлено у відповідність дійсне число $P(A)$, $A \in \mathcal{A}$, яке задовольняє такі умови:

а) $P(\Omega) = 1$;

б) $P(A) \geq 0$;

в) якщо $A_1 \in \mathcal{A}, A_2 \in \mathcal{A}, \dots, A_n \in \mathcal{A}$ і множини $A_1, A_2, \dots, A_n$ не перетинаються (тобто події є несумісними): $A_i \cap A_j = \emptyset, (i \neq j)$, то

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

Функція $P(\cdot)$ називається **скінченно-адитивною ймовірнісною мірою** на алгебрі.

2) Нехай задана алгебра $\mathcal{A}$ і кожному елементу алгебри ставиться у відповідність число $P(A)$, $A \in \mathcal{A}$, яке задовольняє такі умови:

а) $P(\Omega) = 1$;

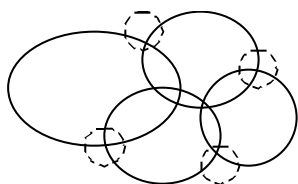
б) $P(A) \geq 0$;

в) якщо $A_1 \in \mathcal{A}$, $A_2 \in \mathcal{A}$, ..., $A_n \in \mathcal{A}$, ..., множини $A_1, A_2, \dots, A_n, \dots$ не перетинаються і, крім того, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, то виконується рівність

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

Така функція $P(\cdot)$ називається **зліченно-адитивною ймовірнісною мірою** на алгебрі.

Теорема Каратеодорі* (про продовження міри). Якщо клас $\mathcal{A}$ – алгебра подій і $P(\cdot)$ – зліченно-адитивна міра на ній, то на мінімальній σ -алгебрі, яка містить алгебру $\mathcal{A}$ ($\sigma(\mathcal{A})$), існує єдина зліченно-адитивна міра $Q(\cdot)$ на $\sigma(\mathcal{A})$, така, що для будь-якої події A з алгебри $\mathcal{A}$ ($A \in \mathcal{A}$) $Q(A) = P(A)$.



Маємо алгебру подій $\mathcal{A}$, на ній міра $P(\cdot)$. Розширюємо $\mathcal{A}$ до мінімальної σ -алгебри. На ній існує єдина міра $Q(\cdot)$, що дорівнює $P(A)$ для будь-якого $A \in \mathcal{A}$.

* Константінос Каратеодорі (Κωνσταντίνος Καραθεοδωρή, 1873-1950) – грецький і німецький математик.

§5. Аксиоматика теорії ймовірності

Ще у кінці XIX сторіччя вчені підійшли до необхідності формулювання системи аксіом теорії ймовірностей. Сформулював аксіоми академік А.М.Колмогоров* у 1933 році.

Нехай розглядається деякий стохастичний експеримент. Відповідний простір елементарних подій позначимо Ω . Нехай $\mathcal{F}$ – клас подій, який задовольняє такі умови:

$A_1)$ Ω як подія належить класу $\mathcal{F}$, тобто $\Omega \in \mathcal{F}$;

$A_2)$ якщо A є подією з класу $\mathcal{F}$, то $\bar{A}$ є подією з класу $\mathcal{F}$:

$$A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F};$$

$A_3)$ якщо A_i ($i = 1, 2, \dots$) є подіями з класу $\mathcal{F}$, то $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

(A_1 – A_3 – умови того, що $\mathcal{F} \in \sigma$ -алгеброю подій). Множини з класу $\mathcal{F}$ називають **випадковими подіями**.

Нехай задана така функція $P(\cdot)$ подій (відображення класу $\mathcal{F}$), яка задовольняє умови:

$B_1)$ $P(\Omega) = 1$;

$B_2)$ $P(A) \geq 0$ для будь-якої події $A \in \mathcal{F}$;

$B_3)$ якщо A_i ($i = 1, 2, \dots$) є елементами класу $\mathcal{F}$, такими, що

$$A_k \cap A_l = \emptyset \quad (k \neq l), \text{ то } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

(умови B_1 – B_3 означають, що $P(\cdot)$ – зліченно-адитивна на одиницю міри на цій структурі). Тоді $P(\cdot)$ називається зліченно-адитивною ймовірнісною мірою на класі $\mathcal{F}$, або якщо випадкова подія $A \in \mathcal{F}$, то $P(A)$ називається **ймовірністю випадкової події A** .

* Андрій Миколайович Колмогоров (1903-1987) – радянський математик.

Перші три аксіоми (A_1-A_3) фактично визначають клас спостережуваних подій (подій, які допускаються в розгляд), останні три аксіоми (B_1-B_3) визначають імовірність цих подій.

Теорія ймовірностей не займається тим, як будувати міру подій. Всюди вважається, що задано $(\Omega, \mathcal{F}, P)$. Тоді кажуть, що задано **ймовірнісний простір**.

Ця система аксіом не суперечлива, бо є об'єкти (наприклад, ті, що розглянуті вище), на яких вони виконуються.

§6. Властивості ймовірності

1⁰. Імовірність протилежної події $P(\bar{A}) = 1 - P(A)$.

◀ Якщо $A \in \mathcal{F}$, то, за аксіомою A_2 , $\bar{A} \in \mathcal{F}$. Але $A \cup \bar{A} = \Omega$, $A \cap \bar{A} = \emptyset$, тому $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$, за аксіомою B_3 . Звідси одержуємо потрібну рівність. ►

Приклад 1. Підкидають гральний кубик. Яка ймовірність того, що не випаде одиниця?

◀ Нехай подія A – при підкиданні кубика випала одиниця. Тоді протилежна до неї подія $\bar{A}$ – при підкиданні кубика не випала одиниця. Тому, згідно з властивістю **1⁰**, $P(\bar{A}) = 1 - P(A) = 1 - \frac{1}{6} = \frac{5}{6}$. ►

Приклад 2. Знайти ймовірність того, що серед п'яти вибраних навмання цифр є хоч одна з трьох даних цифр?

◀ Нехай подія A – серед п'яти вибраних цифр є хоч одна з трьох даних цифр. Знайдемо ймовірність протилежної до неї події $\bar{A}$ – серед п'яти вибраних цифр немає жодної з трьох даних цифр. Оскільки вибрати п'ять цифр можна 10^5 способами, а вибрати п'ять цифр, які не містять жодної з трьох даних цифр, можна 7^5 способами, то $P(\bar{A}) = \left(\frac{7}{10}\right)^5$. Тому, з властивості **1⁰**, $P(A) = 1 - P(\bar{A}) = 1 - (0,7)^5$. ►

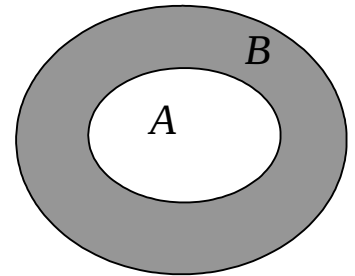
2⁰. Імовірність неможливої події $P(\emptyset) = 0$.

◀ Якщо $A = \Omega$, тоді $\bar{A} = \emptyset$ і з властивості **1⁰**, при $A = \Omega$, одержуємо $P(\emptyset) = 1 - 1 = 0$. ►

3⁰. Нехай події A і B з σ -алгебри $\mathcal{F}$ ($A \in \mathcal{F}$, $B \in \mathcal{F}$), крім цього, подія A викликає подію B ($A \subset B$), тоді ймовірність різниці подій дорівнює різниці їх імовірностей

$$P(B \setminus A) = P(B) - P(A).$$

◀ Зрозуміло, що $B = A \cup (B \setminus A)$, важливо тут, що $A \cap (B \setminus A)$ – це неможлива подія, бо $A \cap (B \setminus A) = A \cap (B \cap \bar{A}) = \emptyset$. Тоді, за аксіомою B_3 , маємо $P(B) = P(A) + P(B \setminus A)$. Звідси одержуємо потрібну рівність. ▶



4⁰. (наслідок з 3⁰). Якщо $A \in \mathcal{F}$, $B \in \mathcal{F}$ і подія A викликає подію B ($A \subset B$), то ймовірність події A не перевищує ймовірності події B :

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

◀ Доведення випливає з властивості 3⁰, бо $P(B \setminus A) \geq 0$ і тоді $P(B) - P(A) \geq 0$. ▶

Приклад 3. Нехай проводиться стохастичний експеримент із підкиданням кубика. Розглянемо події A – на верхній грані кубика випала одиниця і B – випало число, яке при діленні на 3 дає остачу 1. Тоді $A = \{1\}$, $B = \{1, 4\}$ і подія A викликає подію B ($A \subset B$). Імовірності цих подій $P(A) = \frac{1}{6}$, $P(B) = \frac{2}{6} = \frac{1}{3}$ і виконується властивість 4⁰

$$P(A) = \frac{1}{6} \leq \frac{1}{3} = P(B).$$

5⁰. Для будь-якої події A з σ -алгебри $\mathcal{F}$ імовірність $P(A) \leq 1$.

◀ Доведення випливає з властивості 4⁰ для події $B = \Omega$. ▶

Якщо врахувати ще аксіому B_2 , то одержуємо

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

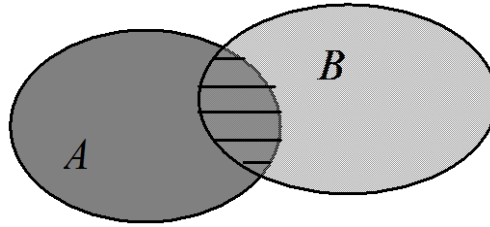
6⁰. (Теорема додавання ймовірностей). Нехай $A \in \mathcal{F}$, $B \in \mathcal{F}$ – випадкові події. Тоді ймовірність суми подій $A \cup B$ дорівнює

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

◀ Запишемо суму подій $A \cup B$ так:

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

У правій частині доданки несумісні, тобто їх добуток є неможливою подією (дорівнює $\emptyset$). Тому ймовірність суми подій,



за аксіомою B_3 , дорівнює сумі ймовірностей доданків. Крім цього, $A \cap B \subset A$ та $A \cap B \subset B$, тому ймовірність різниці (у першому і другому доданках) дорівнює різниці ймовірностей. Тобто

$$\begin{aligned} P(A \cup B) &= P(A \setminus (A \cap B)) + P(B \setminus (A \cap B)) + P(A \cap B) = \\ &= P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B). \end{aligned} \quad \blacktriangleright$$

Приклад 4. Підкидається тетраедр, грані якого замальовані (одна грань – жовта, друга – синя, третя – червона, на четверту – нанесені всі три кольори), і відстежується подія, яким кольором замальована грань тетраедра, якою він впаде на стіл. Яка ймовірність того, що на цій грані є жовтий або синій колір?

◀ Розглянемо такі події:

$\mathcal{J}$ – на тій грані, що випала на стіл, є жовтий колір;

$\mathcal{C}$ – на тій грані, що випала на стіл, є синій колір;

$\mathcal{Ч}$ – на тій грані, що випала на стіл, є червоний колір.

Потрібно знайти $P(\mathcal{K} \cup C)$.

Імовірність випадання тетраедра якоюсь гранню (конкретною) на стіл, у класичному розумінні, дорівнює $\frac{1}{4}$. Підрахуємо ймовірності

$$P(\mathcal{K}) = \frac{2}{4} = \frac{1}{2} = P(C), \quad P(\mathcal{K} \cap C) = \frac{1}{4}. \text{ Тоді, за властивістю } \mathbf{6^0},$$

$$P(\mathcal{K} \cup C) = P(\mathcal{K}) + P(C) - P(\mathcal{K} \cap C) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}. \quad \blacktriangleright$$

$\mathbf{6^0_1}$. (Узагальнення $\mathbf{6^0}$). Легко отримати такий результат:

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

◀ Доведення випливає з властивості $\mathbf{6^0}$ для подій $A \cup B$ та C . ▶

У загальному ймовірність суми n подій дорівнює

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}) + \\ + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right).$$

Ця формула має назву **теорема додавання ймовірностей**, її можна довести за індукцією.

$\mathbf{7^0}$. УМОВИ

а) якщо $A_i \in \mathcal{F}$, $(i=1,2,\dots)$, $A_k \cap A_l = \emptyset$, $(k \neq l)$, то

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

б) якщо $A_i \in \mathcal{F}$, $(i=1,2,\dots)$, $A_1 \subset A_2 \subset A_3 \subset \dots$, то

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} P(A_i);$$

в) якщо $A_i \in \mathcal{F}$, $(i=1,2,\dots)$, $A_1 \supset A_2 \supset A_3 \supset \dots$, то

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} P(A_i);$$

г) якщо $A_i \in \mathcal{F}$, $(i=1,2,\dots)$, $A_1 \supset A_2 \supset A_3 \supset \dots$ і $\bigcap_{i=1}^{\infty} A_i = \emptyset$, то

$$\lim_{i \rightarrow \infty} P(A_i) = 0$$

еквівалентні.

Властивість 7^0 означає, що аксіому B_3 можна замінити на умови б) або в) або г). Ці умови означають:

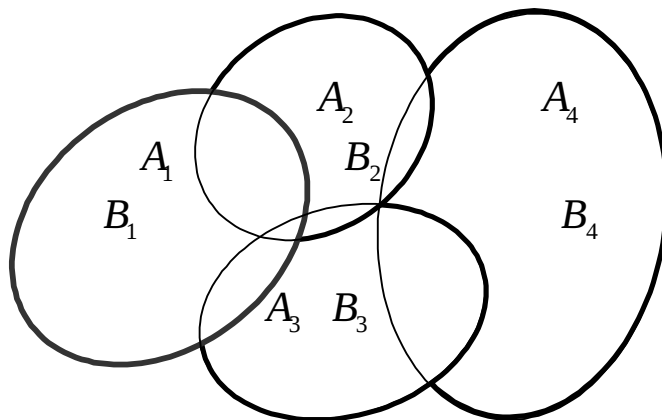
г) імовірність неперервна у нулі (у неможливій події);

б) , в) можливість здійснення „переходу до границі під знаком імовірності”, що можна записати так: $P\left(\lim_{i \rightarrow \infty} A_i\right) = \lim_{i \rightarrow \infty} P(A_i)$.

8^0 . Якщо A_i – елементи σ -алгебри $\mathcal{F}$ ($i=1,2,\dots$), то ймовірність суми подій не перевищує суми їх імовірностей: $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$.

◀ Розглянемо події

$$B_1 = A_1, \quad B_2 = A_2 \cap \bar{A}_1, \quad B_3 = A_3 \cap \bar{A}_1 \cap \bar{A}_2, \dots, B_n = A_n \cap \bar{A}_1 \cap \dots \cap \bar{A}_{n-1}, \dots$$



тобто утворимо множини $B_1, B_2, \dots$, які між собою не перетинаються.

Важливо, що $\bigcup_{i=1}^{\infty} A_i = \bigcup_{n=1}^{\infty} B_n$ і $B_k \cap B_l = \emptyset$, $(k \neq l)$, крім цього,

$B_n \subset A_n$, тому $P(B_n) \leq P(A_n)$. Тоді

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n) \leq \sum_{i=1}^{\infty} P(A_i). \quad \blacktriangleright$$

9⁰. Якщо A_i – елементи σ -алгебри $\mathcal{F}$ ($i=1,2,\dots$), то ймовірність добутку подій $P\left(\bigcap_{i=1}^{\infty} A_i\right) \geq 1 - \sum_{i=1}^{\infty} P(\bar{A}_i)$.

$$\blacktriangleleft P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1 - P\left(\overline{\bigcap_{i=1}^{\infty} A_i}\right) \overset{\text{за законом де Моргана}}{=} 1 - P\left(\bigcup_{i=1}^{\infty} \bar{A}_i\right) \overset{\text{за 8⁰}}{\geq} 1 - \sum_{i=1}^{\infty} P(\bar{A}_i). \quad \blacktriangleright$$

Зауваження 1. Теорема додавання ймовірностей у випадку, коли події несумісні, формулюється так: ймовірність суми подій дорівнює сумі їх ймовірностей:

$$P(A \cup B) = P(A) + P(B) \quad \text{при} \quad A \cap B = \emptyset.$$

Зауваження 2. Якщо ймовірність події дорівнює нулю, то це не означає, що подія неможлива.

Наприклад, на відрізку $[0, 1]$ вибираємо число. Розглянемо подію: „вибране число раціональне”. Подія ця можлива, але міра відповідної множини дорівнює нулю і тому ймовірність події дорівнює нулю.

Але якщо подія неможлива, то її ймовірність дорівнює нулю.

Отже, рівність $P(A \cup B) = P(A) + P(B)$ виконується не лише у випадку $A \cap B = \emptyset$, але й коли ймовірність можливої події дорівнює нулю (це впливає з попереднього зауваження 2).

§7. Умовні ймовірності

Нехай задано ймовірнісний простір $(\Omega, \mathcal{F}, P)$. Припустимо, що деяка подія $B \in \mathcal{F}$ виконалася. Якщо задана ще деяка подія $A \in \mathcal{F}$, то природно виникає питання: чи залежить ймовірність події A від того, що відомо, чи подія B виконалася, чи невідомо нічого про подію B ?

Приклад. Підкидаються два гральних кубики. Простір елементарних подій Ω у цьому експерименті складається з 36 елементарних подій $\Omega = \{(1,1), (1,2), \dots, (1,6), \dots, (6,6)\}$. $\mathcal{F}$ – це множина всіх підмножин множини Ω . Нехай подія $B = \{\text{хоча б на одному з гральних кубиків випала одиниця}\}$, тобто

$$B = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (3,1), (4,1), (5,1), (6,1)\}.$$

Нехай подія $A = \{\text{сума очок на двох кубиках не перевищує чотирьох}\}$. Тобто $A = \{(1,1), (1,2), (1,3), (2,1), (2,2), (3,1)\}$.

Обчислимо ймовірність події A : $P(A) = 6/36 = 1/6$. Припустимо, що надійшла інформація про те, що подія B відбулася. Знайдемо ймовірність $P_{B \text{ відбулася}}(A)$ події A , за умови, що подія B відбулася.

Простором елементарних подій тепер є не Ω , а множина B , бо відомо, що подія B відбулася. І загальна кількість елементарних подій дорівнює 11. У множині A є 5 сприятливих наслідків. Тому

$$P_{B \text{ відбулася}}(A) = 5/11.$$

Знайдемо ще $P(A \cap B)$. Спільних елементів в A і B 5, а весь простір Ω містить 36 елементарних наслідків, тобто, за класичним означенням ймовірності, $P(A \cap B) = 5/36$. Зауважимо, що $P(B) = 11/36$. Проаналізувавши три останні ймовірності, приходимо до висновку, що у цьому прикладі

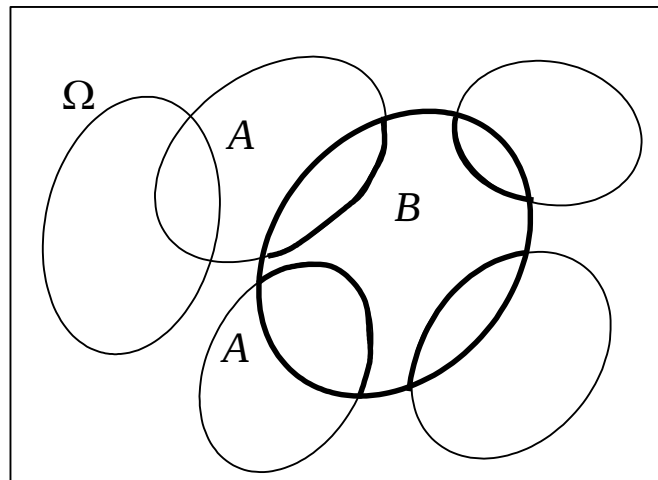
$$P_{B \text{ відбулася}}(A) = \frac{P(A \cap B)}{P(B)} = \frac{5/36}{11/36} = 5/11.$$

Узагальнимо цей приклад.

Означення 1. Нехай B – така подія, що $P(B) > 0$. Тоді **умовною ймовірністю** події A , за умови, що подія B відбулася (позначається $P_B(A)$ або $P(A|B)$), називається відношення ймовірності добутку подій A і B до ймовірності події B :

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

Умовна ймовірність $P(A|B)$ є ймовірністю на деякому іншому ймовірнісному просторі. Дійсно, нехай маємо деякий ймовірнісний простір $(\Omega, \mathcal{F}, P)$. Замість Ω перейдемо до іншого простору елементарних подій – B , отже $\Omega \sim B$ ($B \in \mathcal{F}$). Нехай A – будь-яка подія з $\mathcal{F}$. Тоді $A \cap B \in \mathcal{F}$, а ще дуже важливо, що $A \cap B \subset B$. Легко



довести, що множини $A \cap B$ утворюють σ -алгебру $\mathcal{F}_B$ підмножин множини B , тобто перетини всіх множин A з множиною B утворюють σ -алгебру.

Отже, σ -алгебрі $\mathcal{F}$ поставили у відповідність σ -алгебру $\mathcal{F}_B$. Легко показати, що функція множини $P(\cdot|B)$ визначає ймовірнісну міру P_B на σ -алгебрі $\mathcal{F}_B$. Тобто маємо ймовірнісний простір

$(B, \mathcal{F}_B, P_B)$. Отже, побудований новий імовірнісний простір $(B, \mathcal{F}_B, P_B)$, у якому ймовірність $P(A|B)$ вже стала безумовною.

Будемо залишатися у ймовірнісному просторі $(\Omega, \mathcal{F}, P)$ і розглядати цю ж імовірність $P(A|B)$, але як умовну.

Властивості умовної ймовірності

$$1^0. P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1. \quad P(B|B) = 1. \quad P(\emptyset|B) = 0.$$

2⁰. $P(A \cap B) = P(B) \cdot P(A|B)$ – це еквівалентно означенню умовної ймовірності. Еквівалентність тут у розумінні, що $P(B) \neq 0$.

3⁰. Якщо $P(A) > 0$ і $P(B) > 0$, то

$$P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$$

– **теорема множення ймовірностей**: імовірність добутку подій дорівнює добутку їх імовірностей, один із співмножників є умовною ймовірністю.

Розглянемо узагальнення теореми множення. Спочатку для трьох подій A, B і C при $P(A \cap B) > 0$ одержуємо:

$$P(A \cap B \cap C) = P(A \cap B) \cdot P(C|A \cap B) = P(A) \cdot P(B|A) \cdot P(C|A \cap B).$$

А тепер для n подій $A_1, A_2, \dots, A_n$ знайдемо ймовірність їх добутку:

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n) &= \\ &= P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{n-1}) \cdot P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}) = \\ &= P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{n-2}) \cdot P(A_{n-1} | A_1 \cap A_2 \cap \dots \cap A_{n-2}) \times \\ &\times P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}) = \dots = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \times \end{aligned}$$

$$\times P(A_4 | A_1 \cap A_2 \cap A_3) \cdot \dots \cdot P(A_n | A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{n-1}).$$

Це формальний запис, бо потрібно, щоб усі ймовірності у правій частині рівності існували. Враховуючи, що

$$A_1 \cap A_2 \cap \dots \cap A_{n-1} \subset A_1 \cap A_2 \cap \dots \cap A_{n-2} \subset \dots \subset A_1 \cap A_2 \subset A_1$$

одержуємо

$$P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \leq P(A_1 \cap A_2 \cap \dots \cap A_{n-2}) \leq \dots \leq P(A_1 \cap A_2) \leq P(A_1).$$

Отже, досить вимагати, щоб $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$, і тоді всі умовні ймовірності у теоремі множення ймовірностей існують.

Означення 2. Події A і B називаються **незалежними**, якщо

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

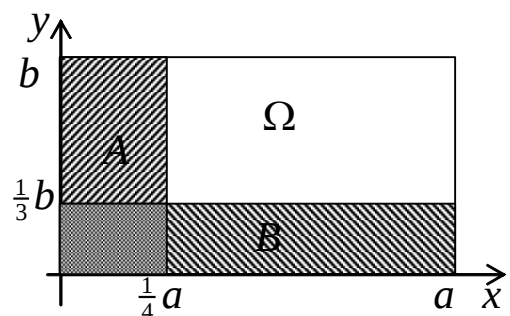
Легко зрозуміти, що події A і B незалежні тоді і тільки тоді, коли $P(A | B) = P(A)$ (якщо відомо, що $P(B) > 0$), або $P(B | A) = P(B)$ (у випадку, коли $P(A) > 0$).

Якщо хоч одна з ймовірностей $P(A)$ або $P(B)$ дорівнює нулю (тобто $P(A) \cdot P(B) = 0$), то $P(A \cap B) = P(A) \cdot P(B)$. Дійсно, $A \cap B \subset B$ і $A \cap B \subset A$. Тоді $0 \leq P(A \cap B) \leq P(A) = 0$ або $0 \leq P(A \cap B) \leq P(B) = 0$, тобто $P(A \cap B) = 0$. Отже, якщо ймовірність хоч однієї з двох подій дорівнює нулю, то ці події незалежні.

Поняття незалежності подій і несумісності подій не тотожні. Події можуть бути незалежні, але сумісні, або несумісні, але залежні.

Приклад. Розглянемо прямокутник, який є простором елементарних подій. Дві частини A і B цього прямокутника мають спільну область, як показано на рисунку. Зрозуміло, що

$$S_\Omega = a \cdot b, \quad S_A = \frac{1}{4} a \cdot b, \quad S_B = \frac{1}{3} a \cdot b,$$



$$S_{A \cap B} = \frac{1}{12} a \cdot b. \quad \text{Тоді} \quad P(A \cap B) = \frac{1}{12}, \quad P(A) = \frac{1}{4}, \quad P(B) = \frac{1}{3},$$

$P(A \cap B) = P(A) \cdot P(B)$. Отже, події A і B сумісні, але незалежні.

Сформулюємо і доведемо основні властивості незалежних подій.

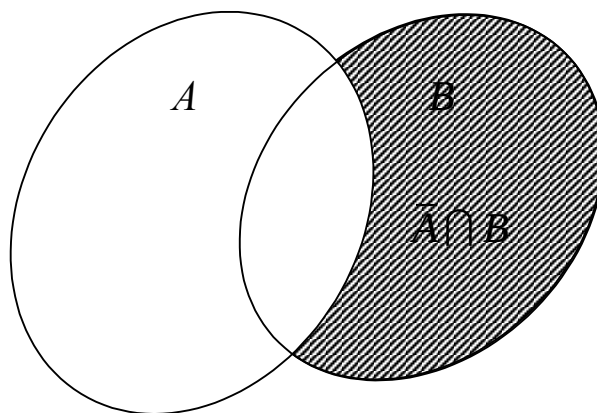
Твердження 1. *Якщо події A і B незалежні, то і події $\bar{A}$ і B – незалежні, події A і $\bar{B}$ – незалежні, події $\bar{A}$ і $\bar{B}$ – незалежні.*

◀ Досить довести, що $\bar{A}$ і B – незалежні. Дано A і B – незалежні, тобто $P(A \cap B) = P(A) \cdot P(B)$. Розглянемо добуток подій $\bar{A}$ і B : $\bar{A} \cap B = B \setminus (A \cap B)$, крім цього, $A \cap B \subset B$. Тоді

$$P(\bar{A} \cap B) = P(B \setminus (A \cap B)) = P(B) - P(A \cap B),$$

але A і B – незалежні, тому

$$P(\bar{A} \cap B) = P(B) - P(A) \cdot P(B) = P(B) \cdot (1 - P(A)) = P(\bar{A}) \cdot P(B).$$



Це й означає, що $\bar{A}$ і B – незалежні.

Те, що події A і $\bar{B}$ – незалежні, доводиться аналогічно. А для доведення того, що події $\bar{A}$ і $\bar{B}$ – незалежні, використовуються два попередні факти. ►

Твердження 2. *Якщо події A і B_1 незалежні, події A і B_2 незалежні, події B_1 і B_2 несумісні (тобто $B_1 \cap B_2 = \emptyset$), тоді події A і $B_1 \cup B_2$ – незалежні.*

◀ Відомо, що $P(A \cap B_1) = P(A) \cdot P(B_1)$, $P(A \cap B_2) = P(A) \cdot P(B_2)$. Розглянемо добуток подій $A \cap (B_1 \cup B_2) = (A \cap B_1) \cup (A \cap B_2)$. Оскільки події B_1 і B_2 несумісні, то

$$(A \cap B_1) \cap (A \cap B_2) = A \cap B_1 \cap B_2 = A \cap \emptyset = \emptyset,$$

тобто події $(A \cap B_1)$ і $(A \cap B_2)$ – також несумісні. Тоді

$$\begin{aligned} P(A \cap (B_1 \cup B_2)) &= P(A \cap B_1) + P(A \cap B_2) = \\ &= P(A) \cdot P(B_1) + P(A) \cdot P(B_2) = P(A) \cdot (P(B_1) + P(B_2)) = \\ &= P(A) \cdot P(B_1 \cup B_2). \end{aligned}$$

B_1, B_2 – несумісні

Отже, події A і $B_1 \cup B_2$ – незалежні. ▶

Умова несумісності подій B_1 та B_2 істотна і має не лише технічний характер для доведення.

Означення 3. Події $A_1, A_2, \dots, A_n$ називаються **попарно незалежними**, якщо для будь-яких i, j , таких, що $1 \leq i \leq n, 1 \leq j \leq n, i \neq j$, імовірність добутку подій A_i та A_j дорівнює добутку їх імовірностей:

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j).$$

Означення 4. Події $A_1, A_2, \dots, A_n$ називаються **незалежними у сукупності**, якщо для будь-якого набору індексів $i_1, \dots, i_k$, таких, що $1 \leq i_1 < \dots < i_k \leq n, 2 \leq k \leq n$, виконується рівність

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j}).$$

Зрозуміло, що з незалежності у сукупності випливає попарна незалежність. Виникає питання: чи не є ці поняття еквівалентними? Тобто чи з попарної незалежності випливає незалежність у сукупності? Відповідь на це питання негативна. З попарної незалежності не

впливає незалежність у сукупності (про це свідчить приклад С.Н.Бернштейна^{*}).

Приклад С.Н.Бернштейна. Розглянемо такий стохастичний експеримент. Підкидається тетраедр, грані якого замальовані (одна грань – жовта, друга – синя, третя – червона, на четверту – нанесені всі три кольори), і відстежується подія, яким кольором замальована грань тетраедра, якою він впаде на стіл. Розглядаються такі події:

$Ж$ – на тій грані, що випала на стіл, є жовтий колір;

$С$ – на тій грані, що випала на стіл, є синій колір;

$Ч$ – на тій грані, що випала на стіл, є червоний колір.

Покажемо, що ці події попарно незалежні, але не є незалежними у сукупності.

Усі грані вважаємо рівноправними. Тому ймовірність випадання тетраедра якоюсь гранню (конкретною) на стіл, у класичному розумінні, дорівнює $\frac{1}{4}$.

Підрахуємо ймовірності подій $Ж$, $С$ і $Ч$:

$$P(Ж) = \frac{2}{4} = \frac{1}{2} = P(С) = P(Ч).$$

Перевіримо попарну незалежність цих подій:

$$P(Ж \cap С) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(Ж) \cdot P(С).$$

Звідси випливає, що події $Ж$ і $С$ – незалежні. Аналогічно,

$$P(Ж \cap Ч) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(Ж) \cdot P(Ч),$$

$$P(С \cap Ч) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(С) \cdot P(Ч).$$

^{*} Сергій Натанович Бернштейн (1880-1968) – радянський математик.

Тобто події $\mathcal{K}$ і $\mathcal{C}$ та $\mathcal{C}$, $\mathcal{C}$ незалежні, а тому події $\mathcal{K}$, $\mathcal{C}$, $\mathcal{C}$ – попарно незалежні. Чи вони незалежні у сукупності? Оскільки

$$P(\mathcal{K} \cap \mathcal{C} \cap \mathcal{C}) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(\mathcal{K}) \cdot P(\mathcal{C}) \cdot P(\mathcal{C}),$$

то події $\mathcal{K}$, $\mathcal{C}$, $\mathcal{C}$ – не є незалежні у сукупності.

Повернемось до твердження 2. Проілюструємо на прикладі С.Н.Бернштейна, що умова несумісності подій B_1 і B_2 у твердженні 2 істотна.

Візьмемо за подію A подію $\mathcal{K}$ з прикладу С.Н.Бернштейна, за подію B_1 – подію $\mathcal{C}$, за подію B_2 – подію $\mathcal{C}$.

Перші дві умови з твердження виконуються: A і B_1 – незалежні та A і B_2 – незалежні (доведено вище). Але тут події B_1 і B_2 сумісні: $B_1 \cap B_2 = \mathcal{C} \cap \mathcal{C} \neq \emptyset$. Обчислимо ймовірність суми подій B_1 і B_2 :

$$P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1 \cap B_2) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}.$$

Тому

$$P(A) \cdot P(B_1 \cup B_2) = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}.$$

Знайдемо ймовірність добутку подій A та $B_1 \cup B_2$:

$$\begin{aligned} P(A \cap (B_1 \cup B_2)) &= P((A \cap B_1) \cup (A \cap B_2)) = \\ &= P(A \cap B_1) + P(A \cap B_2) - P(A \cap B_1 \cap B_2) = \frac{1}{4} + \frac{1}{4} - \frac{1}{4} = \frac{1}{4}. \end{aligned}$$

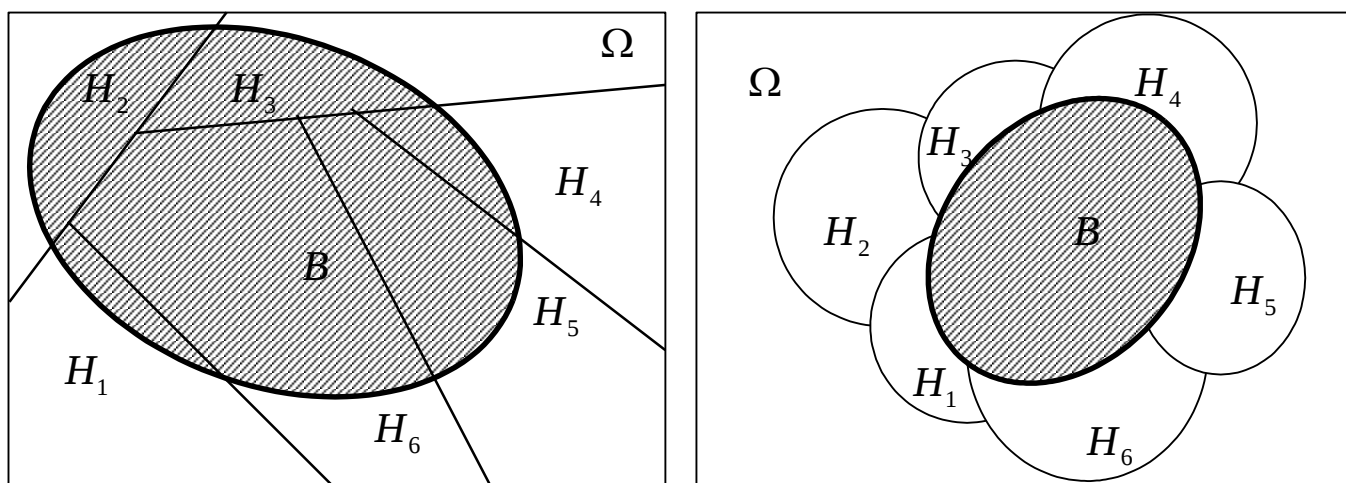
Тобто

$$P(A \cap (B_1 \cup B_2)) = \frac{1}{4} \neq \frac{3}{8} = P(A) \cdot P(B_1 \cup B_2).$$

Звідси випливає, що умова несумісності подій B_1 і B_2 у твердженні 2 істотна.

§8. Формули повної ймовірності, Бейєса

Нехай задано ймовірнісний простір $(\Omega, \mathcal{F}, P)$. Проводиться стохастичний експеримент, у якому спостерігається деяка подія B з $\mathcal{F}$. Нехай також задано деякі події $H_1, H_2, \dots, H_n$, $H_i \in \mathcal{F}$, які є несумісними, тобто $H_i \cap H_j = \emptyset$, $i \neq j$. Якщо B викликає суму подій $\bigcup_{i=1}^n H_i$, тобто $B \subset \bigcup_{i=1}^n H_i$, то кажуть, що події $H_1, H_2, \dots, H_n$ є **гіпотезами** відносно події B (складають систему гіпотез відносно події B). Якщо $\bigcup_{i=1}^n H_i = \Omega$, то кажуть, що **система гіпотез повна**.



На лівому рисунку система гіпотез повна.

Припустимо, що задана система гіпотез відносно події B і ймовірності $P(H_i) > 0$, $i = \overline{1, n}$. Подію B можна записати так:

$$B = B \cap \left(\bigcup_{i=1}^n H_i \right) = \bigcup_{i=1}^n (B \cap H_i)$$

(тут використано розподільчу властивість імовірності).

Оскільки H_i неперетинні (несумісні події), то $(B \cap H_i)$ та $(B \cap H_j)$ ($i \neq j$) також несумісні, тобто: $(B \cap H_i) \cap (B \cap H_j) = \emptyset$,

($i \neq j$). Тому ймовірність суми подій дорівнює сумі ймовірностей цих подій:

$$P(B) = P\left(\bigcup_{i=1}^n (B \cap H_i)\right) = \sum_{i=1}^n P(B \cap H_i), \quad (1)$$

але

$$P(B \cap H_i) = P(H_i) \cdot P(B | H_i). \quad (2)$$

Підставляючи рівності (2) в (1), одержуємо:

$$P(B) = \sum_{i=1}^n P(H_i) \cdot P(B | H_i)$$

– **формулу повної ймовірності.**

Припустимо тепер додатково, що $P(B) > 0$. Тоді, взагалі кажучи, можна було б записати:

$$P(B \cap H_i) = P(B) \cdot P(H_i | B). \quad (3)$$

Отже, прирівнюючи (2) і (3), можна записати такі рівності:

$$P(H_i) \cdot P(B | H_i) = P(B) \cdot P(H_i | B), \quad i = \overline{1, n}.$$

Звідси визначаємо $P(H_i | B)$:

$$P(H_i | B) = \frac{P(H_i) \cdot P(B | H_i)}{P(B)}, \quad i = \overline{1, n},$$

але у знаменнику можна використати формулу повної ймовірності. Тоді одержимо:

$$P(H_i | B) = \frac{P(H_i) \cdot P(B | H_i)}{\sum_{k=1}^n P(H_k) \cdot P(B | H_k)}, \quad i = \overline{1, n}$$

– **формули Бейєса*.**

* *Томас Бейєс (Thomas Bayes) (1702-1761) – англійський математик.*

Суть цих формул така: за формулою повної ймовірності – можна обчислити повну (безумовну) ймовірність через умовні ймовірності; формула Бейєса – якщо відомо, що якась подія відбулася, то така додаткова інформація дає змогу перерахувати (переоцінити) гіпотези й отримати ймовірності гіпотез, при умові, що деяка подія відбулася.

Підкреслимо, що $\sum_{i=1}^n P(H_i | B) = 1$. Формально це одержати просто: $P(B) = \sum_{i=1}^n P(H_i) \cdot P(B | H_i)$ і, поділивши на $P(B)$, матимемо, що $\sum_{i=1}^n P(H_i | B) = 1$. Хоча, важливо підкреслити, що $\sum_{i=1}^n P(H_i) \leq 1$, причому тут досягається рівність лише у випадку повної системи гіпотез, тобто при $\bigcup_{i=1}^n H_i = \Omega$, і ліва частина буде строго меншою одиниці при $\bigcup_{i=1}^n H_i \subset \Omega$.

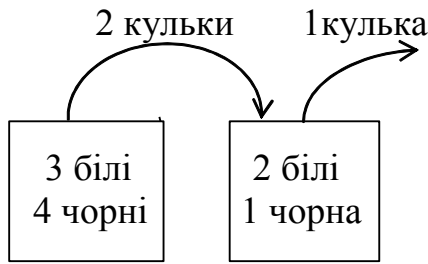
Приклад. У двох ящиках є білі і чорні кульки. У першому з них три білих і чотири чорних кульки, у другому – дві білих і одна чорна. З першого ящика навмання перекладають дві кульки у другий, після цього з другого виймають одну кульку. Ставляться питання: а) яка ймовірність того, що вийнята біла кулька? б) яка ймовірність того, що з першого ящика в другий перекладено дві білих кульки, якщо вийнята з другого ящика кулька виявилася білою?

◀ Позначимо подію B – з другого ящика вийнята біла кулька. Висунемо такі гіпотези:

H_1 – з першого ящика у другий переклали дві білі кульки;

H_2 – з першого ящика у другий переклали одну білу і одну чорну кульки;

H_3 – з першого ящика у другий переклали дві чорні кульки.



Ці гіпотези несумісні і утворюють повну групу подій (гіпотез). Враховуючи прийняті позначення, у пункті а) треба обчислити $P(B)$, а у пункті б) $P(H_1 | B)$.

$$P(H_1) = \frac{C_3^2}{C_7^2} = \frac{3 \cdot 2}{7 \cdot 6} = \frac{1}{7}, \quad P(H_2) = \frac{C_3^1 \cdot C_4^1}{C_7^2} = \frac{2 \cdot 3 \cdot 4}{7 \cdot 6} = \frac{4}{7},$$

$$P(H_3) = \frac{C_4^2}{C_7^2} = \frac{4 \cdot 3}{7 \cdot 6} = \frac{2}{7}, \quad P(B | H_1) = \frac{4}{5}, \quad P(B | H_2) = \frac{3}{5}, \quad P(B | H_3) = \frac{2}{5},$$

$$P(B) = \frac{1}{7} \cdot \frac{4}{5} + \frac{4}{7} \cdot \frac{3}{5} + \frac{2}{7} \cdot \frac{2}{5} = \frac{20}{35} = \frac{4}{7}, \quad P(H_1 | B) = \frac{1}{5}.$$

Зауважимо, що ймовірність першої гіпотези дорівнювала $P(H_1) = \frac{1}{7}$, а у результаті експерименту, коли стало відомо, що з другого ящика вийняли білу кульку, переоцінили її й одержали $P(H_1 | B) = \frac{1}{5}$.

Задачу можна розв'язати, висунувши й іншу систему гіпотез:

H_1 – кулька, яку вийняли з другого ящика, спочатку знаходилась у першому ящику;

H_2 – кулька, яку вийняли з другого ящика, спочатку знаходилась у другому ящику.

Ці гіпотези також несумісні й утворюють повну групу подій. Тоді

$$P(H_1) = \frac{2}{5}, \quad P(H_2) = \frac{3}{5}, \quad P(B | H_1) = \frac{3}{7}, \quad P(B | H_2) = \frac{2}{3},$$

$$P(B) = \frac{2}{5} \cdot \frac{3}{7} + \frac{3}{5} \cdot \frac{2}{3} = \frac{4}{7}. \quad \blacktriangleright$$

§9. Схема незалежних випробувань. Закон Бернуллі (біноміальний закон)

Припустимо, що один і той самий експеримент проводиться багато разів в одних і тих самих умовах. Результати попередніх експериментів не впливають на результати наступних і, природно, що результати наступних експериментів не впливають на результати попередніх (в цьому полягає незалежність випробувань). У кожному експерименті спостерігається деяка подія A (ведеться спостереження за подією A). Імовірність появи цієї події у кожному з експериментів відома і дорівнює $P(A) = p$. Ставиться питання про те, яка ймовірність того, що в серії експериментів подія A відбувається (буде спостерігатися) певну кількість разів?

Будемо вважати, що проводиться серія з n експериментів. Поставимо питання: яка ймовірність того, що в серії з n експериментів подія A відбудеться m разів, де, зрозуміло, $0 \leq m \leq n$? Імовірність того, що подія A відбудеться в m експериментах, позначимо через $P_n(m)$.

Простір елементарних подій Ω можна подати у вигляді

$$\Omega = \left\{ \underbrace{AAAAA \dots A}_n, \underbrace{A\bar{A}AAA \dots A}_n, \underbrace{AA\bar{A}A\bar{A} \dots A}_n, \underbrace{\bar{A}AAAA \dots A}_n, \dots, \underbrace{\bar{A}\bar{A}\bar{A}\bar{A} \dots \bar{A}}_n \right\}.$$

Всього точок у просторі $\Omega \in 2^n$, бо є n позицій, а на кожній позиції може бути A або $\bar{A}$. Цим точкам треба поставити у відповідність ймовірності. Позначимо $P(\bar{A}) = 1 - P(A) = 1 - p = q$. Тоді, оскільки всі експерименти незалежні, то кожній точці (події) ставляться у відповідність добутки ймовірностей: точці $AAAAA \dots A$ треба поставити у відповідність добуток $ppppp \dots p$, точці $AA\bar{A}A\bar{A} \dots A$ треба поставити у відповідність добуток $ppqrpq \dots p$, точці $\bar{A}\bar{A}\bar{A}\bar{A} \dots \bar{A}$, відповідно, $qqqqq \dots q$ і т.д. Легко переконатись, що сума поставлених

у відповідність імовірностей дорівнює одиниці, дійсно, ця сума фактично дорівнює $\sum_{m=0}^n C_n^m p^m q^{n-m} = (p+q)^n = 1$.

Виділимо ті елементарні результати експерименту, у яких подія A відбулася m разів, тобто ті точки (події) з множини Ω , у яких m місць займає подія A , а решту $n-m$ місць займає подія $\bar{A}$. Відповідні таким подіям (точкам) імовірності дорівнюють $p^m q^{n-m}$. І таких точок є C_n^m , тобто одержуємо:

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

– **формулу Бернуллі*** для $m = 0, 1, 2, \dots, n$.

За цією формулою, очевидно, можна підраховувати й інші ймовірності. Наприклад: яка ймовірність того, що в п'яти експериментах подія відбудеться не менше двох разів (позначається $P_5(\geq 2)$)?

Будемо вважати, що $p = \frac{2}{3}$. Тоді шукана ймовірність

$$P_5(\geq 2) = P_5(2) + P_5(3) + P_5(4) + P_5(5),$$

але її можна визначити простіше, перейшовши до протилежної події

$$\begin{aligned} P_5(\geq 2) &= 1 - P_5(< 2) = 1 - (P_5(0) + P_5(1)) = \\ &= 1 - \left(C_5^0 \cdot \left(\frac{2}{3}\right)^0 \cdot \left(\frac{1}{3}\right)^5 + C_5^1 \cdot \left(\frac{2}{3}\right)^1 \cdot \left(\frac{1}{3}\right)^4 \right) = 1 - 11 \cdot \left(\frac{1}{3}\right)^5. \end{aligned}$$

Можна знаходити і такі ймовірності: подія відбудеться не менше двох і не більше чотирьох разів, просумувавши відповідні ймовірності, знайдені за формулою Бернуллі.

Практично використовувати формулу Бернуллі можна при досить малих значеннях n (через складність обчислень).

* *Якоб Бернуллі (Jakob Bernoulli, 1655-1705) – швейцарський математик.*

п. 9.1. Найбільш імовірне число

Уявімо, що проводиться серія експериментів, у кожному з яких певна подія може відбутися певну кількість разів. Імовірність $P_n(m)$ обчислюється за формулою Бернуллі. Виникає питання: для якого значення m найбільшою є ймовірність $P_n(m)$. Фактично необхідно дослідити функцію $P_n(m)$ як функцію аргументу m : цікавить, при якому m значення функції $P_n(m)$ буде найбільшим. Таке m називається **найбільш імовірним числом**.

Для знаходження точки максимуму функції методи з використанням похідних не можна використати, бо функція не є неперервна. Тому розглянемо відношення:

$$\frac{P_n(m+1)}{P_n(m)} = \frac{C_n^{m+1} p^{m+1} q^{n-m-1}}{C_n^m p^m q^{n-m}} = \frac{(n-m) \cdot p}{(m+1) \cdot q}.$$

Дослідимо, при яких m це відношення більше одиниці (тоді функція $P_n(m)$ зростає), а коли відношення менше одиниці (тоді функція $P_n(m)$ спадає). Одержуємо:

1) $\frac{n-m}{m+1} \cdot \frac{p}{q} > 1$, звідси $np - mp > mq + q$, тому $np - q > m(q + p)$ і

$m < np - q$. Тобто, якщо $m < np - q$, то $P_n(m+1) > P_n(m)$, а тому функція $P_n(m)$ зростає;

2) аналогічно, $\frac{n-m}{m+1} \cdot \frac{p}{q} < 1$, тоді $np - mp < mq + q$, звідси

$np - q < m(q + p)$ і $m > np - q$. Тобто, якщо $m > np - q$, то $P_n(m+1) < P_n(m)$, а тому функція $P_n(m)$ спадає;

3) $\frac{n-m}{m+1} \cdot \frac{p}{q} = 1$, звідси $m = np - q$, тобто ймовірність набуває однако-

ві значення при m і $m+1$: $P_n(m+1) = P_n(m)$.

Але m може бути тільки цілим числом. Отже, висновок такий:
якщо $m = np - q$ – ціле число, то існують два найбільш імовірні числа:

$$m'_0 = np - q \text{ та } m''_0 = m'_0 + 1 = np + p$$

(бо $np - q + 1 = np + (1 - q) = np + p$);

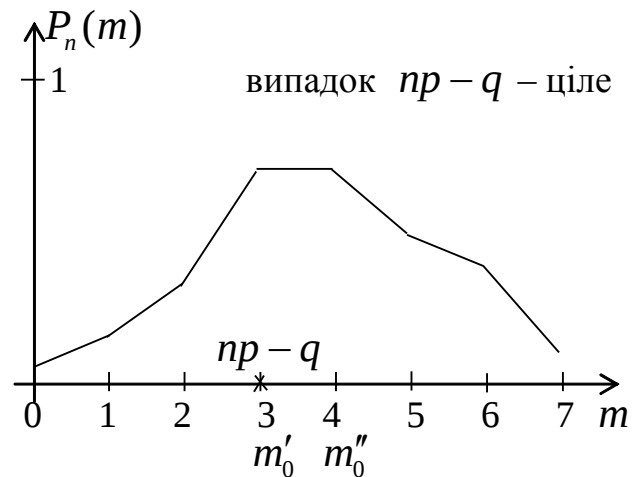
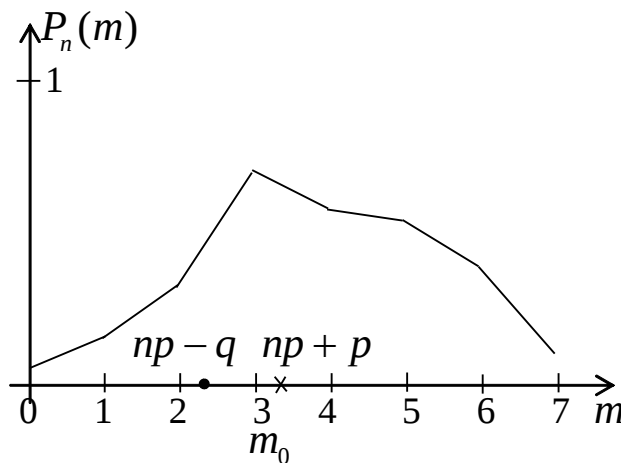
якщо $m = np - q$ – не ціле число, то існує одне найбільш імовірне число, яке міститься між числами $np - q$ та $np + p$, тобто його можна знайти за такою формулою

$$np - q \leq m_0 < np + p.$$

Отже,

$$m_0 = [np - q] + 1, \text{ або } m_0 = [np + p],$$

тут $[.]$ – ціла частина.



Алгоритм знаходження найбільш імовірного числа:

- 1) обчислити np ; якщо np – ціле, то $m_0 = np$;
- 2) якщо np – не ціле, то обчислити $np - q$;
- 3) якщо $np - q$ – не ціле, то найбільш імовірне число

$$m_0 = [np - q] + 1;$$

- 4) якщо $np - q$ – ціле, то є два найбільш імовірних числа

$$m'_0 = np - q \text{ і } m''_0 = m'_0 + 1.$$

п. 9.2. Біноміальний і геометричний розподіли

Нехай розглядається схема незалежних випробувань. Набір чисел $\{P_n(k), k = 0, 1, \dots, n\}$ називається **біноміальним законом розподілу ймовірностей** або **біноміальним розподілом ймовірностей**.

Зі схемою незалежних випробувань тісно пов'язаний ще один розподіл ймовірностей, який називається геометричним розподілом ймовірностей.

Розглянемо такий стохастичний експеримент. Нехай нас цікавить певна подія A , яка в кожному експерименті з'являється з ймовірністю p . Серію експериментів будемо проводити до першої появи події A . Позначимо через $B(m)$ – подію, яка полягає у тому, що до появи події A необхідно провести m експериментів. Простір елементарних подій тут такий:

$$\Omega = \left\{ A, \bar{A}A, \bar{A}\bar{A}A, \bar{A}\bar{A}\bar{A}A, \dots, \underbrace{\bar{A}\bar{A}\dots\bar{A}}_m A, \dots \right\}.$$

Зрозуміло, що ймовірність події $B(m)$ $P(B(m)) = q^m p$, $m = 0, 1, 2, \dots$

Набір чисел $\{q^m p, m = 0, 1, 2, \dots\}$ називається **геометричним розподілом ймовірностей**.

Зауважимо, що сума цих чисел дорівнює одиниці. Такий розподіл ймовірності володіє властивістю відсутності післядії.

Позначимо через $B(\geq k)$ – подію, яка полягає у тому, що до першої появи події A необхідно провести не менше k експериментів. Запишемо таку властивість:

$$P(B(m+n) \mid B(\geq n)) = P(B(m)).$$

Ця властивість і має назву **властивості відсутності післядії** у геометричному розподілі. Переконаємося у правильності останньої формули.

◀ За означенням умовної ймовірності,

$$P(B(m+n) \mid B(\geq n)) = \frac{P(B(n+m) \cap B(\geq n))}{P(B(\geq n))},$$

але $B(\geq n) \supset B(n+m)$, тому $B(n+m) \cap B(\geq n) = B(n+m)$. Крім цього,

$$B(\geq n) = B(n) \cup B(n+1) \cup B(n+2) \cup \dots$$

і події, які записані в останній рівності справа, – несумісні, тому ймовірність суми цих подій дорівнює сумі їх ймовірностей:

$$P(B(\geq n)) = \sum_{i=n}^{\infty} P(B(i)) = \sum_{i=n}^{\infty} p \cdot q^i = q^n \cdot p \cdot \sum_{i=0}^{\infty} q^i = q^n \cdot p \cdot \frac{1}{1-q} = q^n.$$

Отже,

$$P(B(m+n) \mid B(\geq n)) = \frac{P(B(n+m))}{P(B(\geq n))} = \frac{q^{n+m} \cdot p}{q^n} = q^m \cdot p = P(B(m)). \quad \blacktriangleright$$

Можна довести, що єдиним розподілом імовірностей, який має властивість відсутності післядії, є геометричний розподіл.

◀ Від супротивного. Нехай деякий інший розподіл має таку властивість. Деякий розподіл – це набір чисел $\{p_0, p_1, p_2, \dots\}$, таких, що $\sum_{i=0}^{\infty} p_i = 1$, $p_i \geq 0$. Властивість відсутності післядії для записаного розподілу означає, що

$$\frac{p_{n+m}}{\sum_{i=n}^{\infty} p_i} = p_m.$$

З останньої рівності при $n=1$ одержуємо: $p_{m+1} = \sum_{i=1}^{\infty} p_i \cdot p_m$, але

сума $\sum_{i=1}^{\infty} p_i = 1 - p_0$. Таким чином, маємо:

$$p_{m+1} = (1 - p_0) p_m.$$

Підставимо в останню рівність $m = 0$, тоді

$$p_1 = (1 - p_0) p_0.$$

Якщо ж вважатимемо $m = 1$, то

$$p_2 = (1 - p_0) p_1 = (1 - p_0)^2 p_0.$$

При $m = 2$

$$p_3 = (1 - p_0) p_2 = (1 - p_0)^3 p_0$$

і т. д. Звідси, узагальнюючи, одержуємо

$$p_m = (1 - p_0)^m p_0.$$

Цікаві лише випадки, коли $0 < p_0 < 1$ і p_0 – це фактично ймовірність появи якоїсь події в експерименті. Нехай вона дорівнює p , тобто $p_0 = p$, і тоді $1 - p_0 = q$. Отже, маємо

$$p_m = (1 - p)^m \cdot p = q^m \cdot p,$$

тобто вибраний закон розподілу ймовірностей дійсно є геометричним розподілом імовірності. ►

§10. Граничні теореми у схемі незалежних випробувань

Формулу Бернуллі не зручно застосовувати при великих n через складність обчислень. Щоб спростити обчислення треба жертвувати точністю обчислень (ймовірність при цьому не змінюється).

п. 10.1. Локальна теорема Муавра – Лапласа

Теорема 1 (локальна теорема Муавра^{*} – Лапласа^{}).**

Нехай у схемі незалежних випробувань використовуються параметри $p, q, p \geq 0, q \geq 0, p + q = 1$. Якщо $x_{n,m} = \frac{m - np}{\sqrt{npq}}$ рівномірно обмежена сталою C , тобто $|x_{n,m}| \leq C$ і границя $\lim_{n \rightarrow \infty} \sqrt{npq} = \infty$, то

$$\lim_{n \rightarrow \infty} \frac{\sqrt{2\pi} \sqrt{npq} P_n(m)}{e^{-\frac{1}{2}x_{n,m}^2}} = 1.$$

У формулюванні теореми замість рівномірної обмеженості $x_{n,m}$ можна було записати: існує стала C , така, що $|x_{n,m}| \leq C$. Умова $\lim_{n \rightarrow \infty} \sqrt{npq} = \infty$ зрозуміла при сталих p і q . Але можливі експерименти, при яких p і q змінюються при зростанні n . Ця теорема правильна і для таких випадків, тому ця умова є в теоремі. Висновок теореми рівносильний тому, що чисельник і знаменник еквівалентні, тобто

$$\sqrt{npq} P_n(m) \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_{n,m}^2}.$$

^{*} Абрахам Де Муавр (Abraham de Moivre, 1667-1754) – англійський математик французького походження.

^{**} П'єр-Сімон Лаплас (Pierre-Simon Laplace, 1749-1827) – французький математик.

◀ Позначимо $a_n = np$, $b_n = npq$. В умовах теореми $m \xrightarrow[n \rightarrow \infty]{} \infty$, і $n - m \xrightarrow[n \rightarrow \infty]{} \infty$, бо при $n \rightarrow \infty$ зростає кількість експериментів m , в яких подія A відбувається, і кількість експериментів $n - m$, в яких подія не спостерігається. Дійсно, з умови теореми можна записати

$$\begin{aligned} m &= x_{n,m} \sqrt{npq} + np = a_n + \sqrt{b_n} \cdot x_{n,m} = \sqrt{b_n} \left(\frac{a_n}{\sqrt{b_n}} + x_{n,m} \right) = \\ &= \sqrt{b_n} \left(\frac{np}{\sqrt{npq}} + x_{n,m} \right) = \sqrt{b_n} \left(\frac{\sqrt{npq}}{q} + x_{n,m} \right) \geq \sqrt{b_n} \left(\sqrt{\frac{np}{q}} - C \right) \xrightarrow[n \rightarrow \infty]{} \infty \end{aligned}$$

(в останній нерівності використано умову $-C \leq x_{n,m} \leq C$). Аналогічно показуємо, що

$$\begin{aligned} n - m &= n - a_n - \sqrt{b_n} x_{n,m} = n(1 - p) - \sqrt{b_n} x_{n,m} = nq - \sqrt{b_n} x_{n,m} = \\ &= \sqrt{b_n} \left(\frac{nq}{\sqrt{npq}} - x_{n,m} \right) \geq \sqrt{b_n} \left(\sqrt{\frac{nq}{p}} - C \right) \xrightarrow[n \rightarrow \infty]{} \infty \end{aligned}$$

Використовуючи формулу Стірлінга^{*},

$$n! \sim \sqrt{2\pi n} \cdot n^n \cdot e^{-n},$$

запишемо

$$\begin{aligned} P_n(m) &= C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m} \sim \\ &\sim \frac{\sqrt{2\pi n} \cdot n^n \cdot e^{-n} \cdot p^m \cdot q^{n-m}}{\sqrt{2\pi m} \cdot m^m \cdot e^{-m} \cdot \sqrt{2\pi(n-m)} \cdot (n-m)^{n-m} \cdot e^{-n+m}} = \\ &= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{n}{m(n-m)}} \cdot \left(\frac{np}{m} \right)^m \cdot \left(\frac{nq}{n-m} \right)^{n-m}. \end{aligned}$$

^{*} Джеймс Стірлінг (James Stirling, 1692-1770) – шотландський математик.

Отже, маємо

$$\sqrt{npq} \cdot P_n(m) \sim \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{n^2 pq}{m(n-m)}} \cdot \left(\frac{np}{m}\right)^m \cdot \left(\frac{nq}{n-m}\right)^{n-m}.$$

Зауважимо, що

$$\frac{m}{n} = p + \frac{\sqrt{b_n}}{n} x_{n,m} = p + x_{n,m} \sqrt{\frac{pq}{n}} \geq p - C \sqrt{\frac{pq}{n}} \sim p \text{ при } n \rightarrow \infty.$$

Крім цього, розглянемо

$$\frac{n-m}{n} = 1 - \frac{m}{n} \sim 1 - p = q.$$

Тому $\frac{n}{n-m} \sim \frac{1}{q}$. Звідси отримуємо $\sqrt{\frac{n^2 pq}{m(n-m)}} \sim 1$ (при $n \rightarrow \infty$). Отже,

$$\sqrt{npq} \cdot P_n(m) \sim \frac{1}{\sqrt{2\pi}} \cdot \left(\frac{np}{m}\right)^m \cdot \left(\frac{nq}{n-m}\right)^{n-m}$$

і

$$\sqrt{2\pi} \cdot \sqrt{npq} \cdot P_n(m) \sim \left(\frac{np}{m}\right)^m \cdot \left(\frac{nq}{n-m}\right)^{n-m}.$$

Залишилось показати, що

$$\left(\frac{np}{m}\right)^m \cdot \left(\frac{nq}{n-m}\right)^{n-m} \xrightarrow{n \rightarrow \infty} e^{-\frac{1}{2}x_{n,m}^2}.$$

Нехай $x_{n,m} = x$. Оскільки

$$m = np + x\sqrt{b_n}, \quad n-m = nq - x\sqrt{b_n},$$

то

$$\begin{aligned}
\left(\frac{np}{m}\right)^m \cdot \left(\frac{nq}{n-m}\right)^{n-m} &= \left(\frac{np}{np+x\sqrt{b_n}}\right)^{np+x\sqrt{b_n}} \cdot \left(\frac{nq}{nq-x\sqrt{b_n}}\right)^{nq-x\sqrt{b_n}} = \\
&= \left(1+x\frac{\sqrt{b_n}}{np}\right)^{-np-x\sqrt{b_n}} \cdot \left(1-x\frac{\sqrt{b_n}}{nq}\right)^{-nq+x\sqrt{b_n}} = \\
&= \left(1+\frac{qx}{\sqrt{b_n}}\right)^{-np-x\sqrt{b_n}} \cdot \left(1-\frac{px}{\sqrt{b_n}}\right)^{-nq+x\sqrt{b_n}} = \theta_n(x).
\end{aligned}$$

Для доведення теореми досить показати, що $\theta_n(x) \xrightarrow[n \rightarrow \infty]{} e^{-\frac{1}{2}x^2}$, або

$$\ln(\theta_n(x)) \xrightarrow[n \rightarrow \infty]{} -\frac{1}{2}x^2$$

рівномірно на кожному скінченному проміжку. Маємо

$$\ln(\theta_n(x)) = (-np - x\sqrt{b_n}) \cdot \ln\left(1 + \frac{qx}{\sqrt{b_n}}\right) + (-nq + x\sqrt{b_n}) \cdot \ln\left(1 - \frac{px}{\sqrt{b_n}}\right) =$$

одержані логарифми розкладаємо в ряд, обмежившись двома доданками

$$\begin{aligned}
&= (-np - x\sqrt{b_n}) \cdot \left(\frac{qx}{\sqrt{b_n}} - \frac{1}{2} \frac{q^2 x^2}{b_n} + O\left(\frac{q^3}{b_n^{3/2}}\right)\right) + \\
&+ (-nq + x\sqrt{b_n}) \cdot \left(-\frac{px}{\sqrt{b_n}} - \frac{1}{2} \frac{p^2 x^2}{b_n} + O\left(\frac{p^3}{b_n^{3/2}}\right)\right) = \\
&= -\frac{npq}{\sqrt{b_n}} x - qx^2 + \frac{1}{2} qx^2 + O\left(\frac{1}{\sqrt{b_n}}\right) + \frac{npq}{\sqrt{b_n}} x - px^2 + \frac{1}{2} px^2 = \\
&= -\frac{1}{2} qx^2 - \frac{1}{2} px^2 = -\frac{x^2}{2} + O\left(\frac{1}{\sqrt{b_n}}\right).
\end{aligned}$$

Звідси і випливає твердження теореми. ►

Тобто доведено таку наближену рівність:

$$P_n(m) \approx \frac{1}{\sqrt{2\pi} \sqrt{npq}} e^{-\frac{1}{2}x_{n,m}^2}.$$

Або, звідси

$$P_n(m) = C_n^m p^m q^{n-m} \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,m}),$$

де

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

– функція, яка затабульована (таблиця 1 у Додатку В).

Отже, **алгоритм обчислення** $P_n(m)$ при великих n такий:

1. Обчислити $x_{n,m}$.
2. Знайти $\varphi(x_{n,m})$ за таблицями.
3. Обчислити $P_n(m)$ за формулою

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,m}).$$

Виникає питання: при яких n можна використовувати теорему? Якщо виписати залишкові члени і провести акуратно всі обчислення, то можна переконатися, що теорема застосовна вже при $\sqrt{npq} \geq 5$ (навіть при $\sqrt{npq} \geq 4$, хоча при цьому похибка складає приблизно 25%).

Доведена теорема називається локальною, бо в ній оцінюється ймовірність $P_n(m)$ того, що саме m разів відбулася подія.

п. 10.2. Інтегральна теорема Муавра – Лапласа (наслідок локальної теореми)

Теорема 2 (інтегральна теорема Муавра – Лапласа). В умовах теореми 1 для будь-яких a, b ($-\infty < a < b < \infty$) імовірність того, що $a \leq x_{n,m} \leq b$ приблизно дорівнює

$$P_n\{a \leq x_{n,m} \leq b\} \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx.$$

$$\blacktriangleleft P\{a \leq x_{n,m} \leq b\} = \sum_{m:a \leq x_{n,m} \leq b} P_n(m) \approx \sum_{m:a \leq x_{n,m} \leq b} \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_{n,m}^2}.$$

Позначимо

$$\Delta x_{n,m} = x_{n,m+1} - x_{n,m} = \frac{m+1-np}{\sqrt{npq}} - \frac{m-np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}.$$

Тому

$$P\{a \leq x_{n,m} \leq b\} = \frac{1}{\sqrt{2\pi}} \sum_{m:a \leq x_{n,m} \leq b} e^{-\frac{1}{2}x_{n,m}^2} \cdot \Delta x_{n,m} \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx,$$

бо це інтегральна сума для функції $e^{-\frac{1}{2}x^2}$. $\blacktriangleright$

Проаналізуємо, як застосовувати теорему 2. Припустимо, що треба підрахувати таку ймовірність: $P_n\{k_1 \leq m \leq k_2\}$. Зрозуміло, що

$$\begin{aligned} P_n\{k_1 \leq m \leq k_2\} &= P_n\{k_1 - np \leq m - np \leq k_2 - np\} = \\ &= P_n\left\{ \frac{k_1 - np}{\sqrt{npq}} \leq x_{n,m} \leq \frac{k_2 - np}{\sqrt{npq}} \right\}. \end{aligned}$$

Позначимо $a = \frac{k_1 - np}{\sqrt{npq}}$, $b = \frac{k_2 - np}{\sqrt{npq}}$, і тепер можна застосовувати теорему 2.

Отже, для знаходження ймовірності $P_n\{k_1 \leq m \leq k_2\}$ потрібно

1) обчислити $a = \frac{k_1 - np}{\sqrt{npq}}$, $b = \frac{k_2 - np}{\sqrt{npq}}$;

2) застосувати теорему 2.

Приклад 1. Нехай проводиться 100 експериментів ($n=100$). Нехай у кожному експерименті ймовірність певної події дорівнює $p=0,5$. Тоді і $q=0,5$. Яка ймовірність того, що в серії зі 100 експериментів ця подія відбудеться 60 разів?

◀ З умови задачі $\sqrt{npq} = 10 \cdot 0,5 = 5$. Тому можна застосувати теорему 1. За умовою задачі, $m = 60$. Обчислимо

$$x_{n,m} = \frac{60 - 100 \cdot 0,5}{\sqrt{100 \cdot 0,5 \cdot 0,5}} = \frac{10}{5} = 2.$$

Тоді за теоремою 1

$$P_{100}(60) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{5} \cdot e^{-2} = \frac{1}{5} \cdot \varphi(2) = \frac{1}{5} \cdot 0,054 = 0,0108. \blacktriangleright$$

Приклад 2. Знайти ймовірність того, що в цьому ж експерименті подія відбудеться не менше 40 разів і не більше 60.

$$\begin{aligned} \blacktriangleleft P\{40 \leq m \leq 60\} &= P\left\{\frac{40 - np}{\sqrt{npq}} \leq \frac{m - np}{\sqrt{npq}} \leq \frac{60 - np}{\sqrt{npq}}\right\} = \\ &= P\left\{-2 \leq \frac{m - np}{\sqrt{npq}} \leq 2\right\} \stackrel{\text{за т.2}}{\approx} \frac{1}{\sqrt{2\pi}} \int_{-2}^2 e^{-\frac{x^2}{2}} dx. \end{aligned}$$

Такий інтеграл обчислюється за таблицями. Таблиці різні, але найчастіше наводяться таблиці значень такої функції (таблиця 2 у Додатку В)

$$\Phi_0(x) \approx \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \quad x \geq 0.$$

Тоді

$$P\{40 \leq m \leq 60\} \approx 2\Phi_0(2) \approx 2 \cdot 0,4772 = 0,9544. \blacktriangleright$$

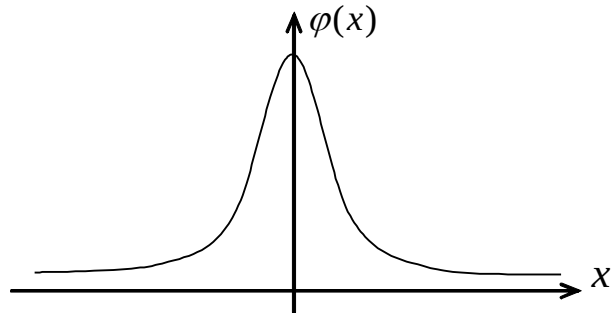
Часом зустрічаються таблиці значень функції

$$\Phi_1(x) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \text{ для } x \leq 0 \text{ або } x \geq 0.$$

Ці всі функції можна просто одержати одна з одної, якщо врахувати, що графік функції

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

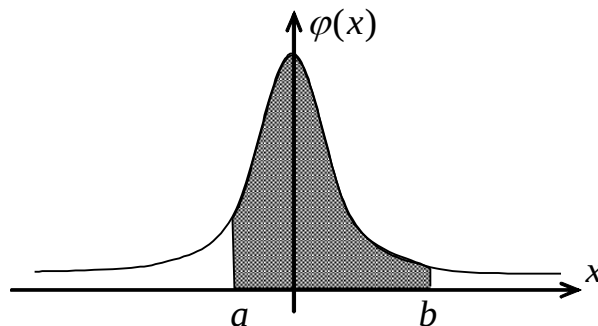
має такий вигляд, як на рисунку,



і те, що інтеграл Пуассона*

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}.$$

Наприклад, при $a < 0$ та $b > 0$ і таблицями значень функції $\Phi_1(x)$ при $x \leq 0$ маємо, що шукана площа дорівнює $1 - \Phi_1(a) - \Phi_1(-b)$.



* Сімеон-Дені Пуассон (Simeon-Denis Poisson) (1781-1840) – французький фізик і математик.

п. 10.3. Теорема Пуассона (теорема про рідкісні події)

Теорема 3 (Пуассона про рідкісні події). Припустимо, що умова теореми 2, а саме $npq \rightarrow \infty$, не виконується, але існує така стала C , що $a_n = np < C$. Тоді при $n \rightarrow \infty$ імовірність $P_n(m)$ дорівнює

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

Назва теореми пов'язана з тим, що тут імовірність p появи події близька до нуля (прямує до нуля). Можливий також випадок, коли ймовірність q неяви події прямує до нуля, але тоді ймовірність появи події p прямує до одиниці, і можна перейти до протилежної події, ймовірність якої вже буде прямувати до нуля.

◀ Схема доведення.

$$\begin{aligned} P_n(m) &= C_n^m p^m q^{n-m} = \frac{n(n-1)(n-2)\dots(n-m+1)}{m!} p^m q^{n-m} = \\ &= \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \dots \cdot \frac{n-m+1}{n} \cdot \frac{n^m}{m!} p^m q^{n-m} = \\ &= 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \dots \cdot \left(1 - \frac{m-1}{n}\right) \cdot \frac{(a_n)^m}{m!} \cdot \left(1 - \frac{np}{n}\right)^n \cdot \left(1 - \frac{np}{n}\right)^{-m} \approx \\ &\approx \frac{(a_n)^m}{m!} \cdot \left(\left(1 - \frac{a_n}{n}\right)^{-\frac{n}{a_n}} \right)^{-a_n} \approx \frac{(a_n)^m}{m!} e^{-a_n}. \quad \blacktriangleright \end{aligned}$$

прямує до 1
при $n \rightarrow \infty$

прямує до e
при $n \rightarrow \infty$

Слід використовувати цю теорему при $p < 0,02$ і досить великих n .

Набір чисел $\left\{ \frac{\lambda^m}{m!} e^{-\lambda}, m = 0, 1, 2, \dots \right\}$ при $\lambda > 0$ утворює розподіл імовірностей, який називається **розподілом Пуассона**.

Фактично теорема стверджує, що при малих p граничним розподілом для біноміального розподілу ймовірності є, при $n \rightarrow \infty$, розподіл Пуассона з параметром $\lambda = a_n = np$.

Для функції $\frac{\lambda^m}{m!} e^{-\lambda}$ складено таблиці для різних значень λ і m (таблиця 3 у Додатку В). Часто затабульована функція $\sum_{m=0}^x \frac{\lambda^m}{m!} e^{-\lambda}$. Використовуючи ці таблиці, можна просто розв'язувати задачі про знаходження $P_\lambda(m)$.

РОЗДІЛ 2

ВИПАДКОВІ ВЕЛИЧИНИ

§1. Поняття випадкової величини

Вище розглядалися такі поняття: стохастичний експеримент, частота, простір елементарних подій, розподіл, випадкова подія. Розглянемо нове поняття випадкової величини, яке буде визначатися при допомозі введеного вже поняття випадкової події.

Випадкова величина – це величина, яка може набувати тих або інших значень у різних випадках. Прикладами випадкових величин можуть бути кількість очок, які випали при одному підкиданні грального кубика, кількість бракованих виробів серед взятих навмання n виробів, кількість попадань у мішень при n пострілах, час безвідмовної роботи приладу, дальність польоту ракети і т.д. Випадкова величина ξ – це число, яке ставиться у відповідність кожному можливому результату експерименту. Оскільки результати експерименту описуються елементарними подіями, випадкову величину можна розглядати як функцію $\xi = \xi(\omega)$ на просторі елементарних подій Ω .

Перш ніж сформулювати означення випадкової величини, розглянемо кілька прикладів.

Приклад 1. Розглянемо такий стохастичний експеримент – двічі підкидають монету. Простір елементарних подій має вигляд $\Omega = \{ГГ, ГР, РГ, РР\}$. Нехай ξ – кількість появ герба. Величина ξ є функцією $\xi = \xi(\omega)$ елементарної події (елементарного результату експерименту). Таблиця значень функції $\xi(\omega)$ має такий вигляд:

ω	$ГГ$	$ГР$	$РГ$	$РР$
$\xi(\omega)$	2	1	1	0

Приклад 2. Підкидають гральний кубик. У цьому випадку простір елементарних подій можна записати у такому вигляді: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, де ω_i означає, що випало i очок. Випадкова величина ξ – кількість очок, що випали – є функцією елементарної події, причому $\xi(\omega) = i$, якщо $\omega = \omega_i$.

Приклад 3. Монету підкидають доти, поки не з'явиться герб. У цьому випадку

$$\Omega = \left\{ Г, РГ, РРГ, РРРГ, \dots, \underbrace{РР\dots Р}_{n-1}Г, \dots \right\} = \{\omega_1, \omega_2, \omega_3, \omega_4, \dots, \omega_n, \dots\}.$$

Нехай ξ – кількість здійснених підкидань монети. Тоді величина ξ є функцією елементарної події, причому $\xi(\omega) = n$, якщо $\omega = \omega_n$ ($n = 1, 2, \dots$).

Ці приклади підтверджують, що випадкові величини можна інтерпретувати як функції на просторі елементарних результатів стохастичного експерименту. Маємо справу з такою функцією $\xi(\omega): \Omega \rightarrow \mathbb{R}$. Це відображення, яке точкам ω з простору елементарних подій Ω ставить у відповідність число. У ймовірнісному просторі $(\Omega, \mathcal{F}, P)$ можна вивчати лише ті події, які належать до $\mathcal{F}$, бо тільки їм поставлено у відповідність числа P . А тій функції $\xi(\omega)$, яка розглядається, можна поставити у відповідність такі події, які є випадковими подіями: $\{\xi(\omega) < 0\}, \dots, \{\xi(\omega) < s\}, \dots$. Ці події можна вивчати, якщо вони належать до σ -алгебри $\mathcal{F}$. Надалі будемо розглядати такі відображення $\xi(\omega)$, що для будь-якого $x \in \mathbb{R}$ подія $\{\xi(\omega) < x\}$ є елементом σ -алгебри $\mathcal{F}$, тобто $\{\xi(\omega) < x\} \in \mathcal{F}$. Таке **відображення** (в теорії міри) називається **вимірним** відносно σ -алгебри $\mathcal{F}$. Часто це записують і так: $\{\omega: \xi(\omega) < x\} \in \mathcal{F}$. Легко зрозуміти, що події $\{\omega: \xi(\omega) > x\}$, $\{\omega: \xi(\omega) = x\}$, $\{\omega: a \leq \xi(\omega) < b\}$ можна представити через подію такого типу $\{\omega: \xi(\omega) < x\}$. Наприклад,

$$\{\omega: a \leq \xi(\omega) < b\} = \{\omega: \xi(\omega) < b\} \setminus \{\omega: \xi(\omega) < a\}.$$

Взагалі, які б події типу $\{\omega: a \leq \xi(\omega) < b\}$ не записали, всі вони належать до σ -алгебри $\mathcal{F}$, якщо тільки $\{\omega: \xi(\omega) < x\} \in \mathcal{F}$, і всі ці події $\{\omega: a \leq \xi(\omega) < b\}$ можна представити зліченим числом подій типу $\{\omega: \xi(\omega) < x\} \in \mathcal{F}$.

Означення 1. Відображення $\xi(\omega): \Omega \rightarrow \mathbb{R}$ називається **випадковою величиною**, якщо для будь-якого дійсного x подія $\{\omega: \xi(\omega) < x\}$ належить до σ -алгебри $\mathcal{F}$.

Можна довести, що у випадку, коли $\xi(\omega)$ і $\eta(\omega)$ дві випадкові величини, то $\xi(\omega) \pm \eta(\omega)$ теж випадкові величини, а також $a \cdot \xi(\omega)$ – випадкова величина. Тому і лінійна комбінація випадкових величин теж випадкова величина. Крім цього, добуток випадкових величин $\xi(\omega) \cdot \eta(\omega)$ є випадковою величиною і частка випадкових величин $\frac{\xi(\omega)}{\eta(\omega)}$ – випадкова величина, якщо ймовірність $P\{\eta(\omega) = 0\} = 0$.

Нехай $\xi(\omega)$ – випадкова величина, тобто для будь-якого $x \in \mathbb{R}$ подія $\{\omega: \xi(\omega) < x\}$ є елементом σ -алгебри $\mathcal{F}$, тоді визначена ймовірність цієї події. Ця ймовірність $P\{\omega: \xi(\omega) < x\} = P\{\xi(\omega) < x\}$ позначається $F_\xi(x)$.

Означення 2. Функція

$$F_\xi(x) = P\{\omega: \xi(\omega) < x\}$$

називається **функцією розподілу випадкової величини $\xi(\omega)$** .

§2. Властивості функції розподілу випадкової величини

1⁰. Функція розподілу $F_\xi(x)$ визначена на всій числовій осі (бо цього вимагає означення).

2⁰. Для будь-якого дійсного x функція розподілу $0 \leq F_\xi(x) \leq 1$ (бо це ймовірність).

3⁰. Функція розподілу $F_\xi(x)$ не спадає.

◀ Треба довести, що з $a < b$ випливає $F_\xi(a) \leq F_\xi(b)$. Розглянемо дві ω -множини $\{\omega: \xi(\omega) < a\}$ і $\{\omega: \xi(\omega) < b\}$. Оскільки $a < b$, то для них правильне таке включення:

$$\{\omega: \xi(\omega) < a\} \subset \{\omega: \xi(\omega) < b\}.$$

Тоді з властивості 4⁰ (монотонності) ймовірності маємо:

$$P\{\omega: \xi(\omega) < a\} \leq P\{\omega: \xi(\omega) < b\}. \quad \blacktriangleright$$

4⁰. $\lim_{x \rightarrow -\infty} F_\xi(x) = 0$, $\lim_{x \rightarrow +\infty} F_\xi(x) = 1$.

◀ У першій частині фактично треба довести, що якщо вибрати довільну послідовність $x \rightarrow -\infty$, то відповідна послідовність $F_\xi(x) \rightarrow 0$. Нехай $x_1 > x_2 > x_3 > \dots > x_n > \dots$, причому $\lim_{n \rightarrow \infty} x_n = -\infty$.

Треба довести, що $\lim_{n \rightarrow \infty} F_\xi(x_n) = 0$. Розглянемо ω -множини: $\{\omega: \xi(\omega) < x_n\}$ і $\{\omega: \xi(\omega) < x_{n+1}\}$. Для них правильне таке включення:

$$\{\omega: \xi(\omega) < x_n\} \supset \{\omega: \xi(\omega) < x_{n+1}\}.$$

Тобто маємо справу зі спадною послідовністю подій, тому

$$\lim_{n \rightarrow \infty} \{\omega: \xi(\omega) < x_n\} = \bigcap_{n=1}^{\infty} \{\omega: \xi(\omega) < x_n\} = \emptyset.$$

Перейдемо до границі в означенні функції розподілу $F_{\xi}(x)$ за властивістю 7⁰в) імовірності і скористаємося властивістю 7⁰г) неперервності ймовірності в нулі (у неможливій події). Одержимо:

$$\lim_{n \rightarrow \infty} F_{\xi}(x_n) = \lim_{n \rightarrow \infty} P\{\xi(\omega) < x_n\} = P\left\{\lim_{n \rightarrow \infty} \{\xi(\omega) < x_n\}\right\} = P(\emptyset) = 0.$$

Аналогічно доводиться друга частина властивості. ►

5⁰. Функція розподілу випадкової величини неперервна зліва, тобто

$$\lim_{x \rightarrow x_0 - 0} F_{\xi}(x) = F_{\xi}(x_0).$$

◀ Розглянемо довільну послідовність $x_1, x_2, \dots, x_n, \dots$, таку, що

$$x_1 < x_2 < x_3 < \dots < x_n < \dots \text{ і } \lim_{n \rightarrow \infty} x_n = x_0.$$

Покажемо, що $\lim_{n \rightarrow \infty} F_{\xi}(x_n) = F_{\xi}(x_0)$.

Запишемо, за означенням, $F_{\xi}(x_n) = P\{\omega : \xi(\omega) < x_n\}$. Порівняємо події (ω -множини) $\{\omega : \xi(\omega) < x_n\}$ і $\{\omega : \xi(\omega) < x_{n+1}\}$. Для них виконується включення $\{\omega : \xi(\omega) < x_n\} \subset \{\omega : \xi(\omega) < x_{n+1}\}$. Отже, послідовність цих подій (множин) монотонно зростає. Тому

$$\lim_{n \rightarrow \infty} \{\omega : \xi(\omega) < x_n\} = \bigcup_{n=1}^{\infty} \{\omega : \xi(\omega) < x_n\} = \{\omega : \xi(\omega) < x_0\}.$$

Тепер можна перейти до границі під знаком імовірності і скористатися властивістю 7⁰б) неперервності ймовірності:

$$\begin{aligned} \lim_{n \rightarrow \infty} F_{\xi}(x_n) &= \lim_{n \rightarrow \infty} P\{\omega : \xi(\omega) < x_n\} = P\left(\lim_{n \rightarrow \infty} \{\omega : \xi(\omega) < x_n\}\right) = \\ &= P\{\omega : \xi(\omega) < x_0\} = F_{\xi}(x_0). \end{aligned} \quad \blacktriangleright$$

Нижче буде наведено багато прикладів, які покажуть, що функція розподілу не обов'язково неперервна справа.

У деяких джерелах в означенні функції розподілу фігурує знак $\leq$. Тоді функція розподілу неперервна і не обов'язково неперервна зліва.

6⁰. Для будь-яких дійсних a і b , пов'язаних співвідношенням $a < b$, має місце рівність

$$P\{\omega : a \leq \xi(\omega) < b\} = F_{\xi}(b) - F_{\xi}(a).$$

◀ Розглянемо дві події (ω -множини): $\{\omega : \xi(\omega) < a\}$ і $\{\omega : \xi(\omega) < b\}$. Зрозуміло, що з умови $a < b$ випливає

$$\{\omega : \xi(\omega) < a\} \subset \{\omega : \xi(\omega) < b\}.$$

Крім цього,

$$\{\omega : a \leq \xi(\omega) < b\} = \{\omega : \xi(\omega) < b\} \setminus \{\omega : \xi(\omega) < a\}.$$

Тому, за властивістю 3⁰ ймовірності, одержуємо

$$P\{\omega : a \leq \xi(\omega) < b\} = P\{\omega : \xi(\omega) < b\} - P\{\omega : \xi(\omega) < a\} = F_{\xi}(b) - F_{\xi}(a). \quad \blacktriangleright$$

7⁰. Імовірність того, що випадкова величина $\xi(\omega)$ набуває значення, яке не перевищує x_0 , дорівнює значенню функції розподілу в точці x_0 справа, тобто

$$P\{\omega : \xi(\omega) \leq x_0\} = F_{\xi}(x_0 + 0).$$

◀ Розглянемо праву частину останньої рівності. За означенням,

$$F_{\xi}(x_0 + 0) = \lim_{x \rightarrow x_0 + 0} F_{\xi}(x) = \lim_{n \rightarrow \infty} F_{\xi}(x_n),$$

де послідовність $\{x_n\}_{n=1}^{\infty}$ така, що $x_1 > x_2 > x_3 > \dots > x_n > \dots$ і $\lim_{n \rightarrow \infty} x_n = x_0$. Розглянемо такі події (ω -множини): $\{\omega : \xi(\omega) < x_n\}$ і $\{\omega : \xi(\omega) < x_{n+1}\}$. Зрозуміло, що $\{\omega : \xi(\omega) < x_n\} \supset \{\omega : \xi(\omega) < x_{n+1}\}$. Тобто послідовність подій $\{\omega : \xi(\omega) < x_n\}$ монотонно спадає, а відповідні множини монотонно звужуються. Тому

$$\lim_{n \rightarrow \infty} \{\omega : \xi(\omega) < x_n\} = \bigcap_{n=1}^{\infty} \{\omega : \xi(\omega) < x_n\} = \{\omega : \xi(\omega) \leq x_0\}.$$

Залишилось перейти до границі за властивістю 7⁰в) імовірності та скористатися властивістю 7⁰г) неперервності ймовірності у неможливій події:

$$P\{\omega: \xi(\omega) \leq x_0\} = P\left\{\bigcap_{n=1}^{\infty} \{\omega: \xi(\omega) < x_n\}\right\} = \lim_{n \rightarrow \infty} P\{\omega: \xi(\omega) < x_n\} = \\ = \lim_{n \rightarrow \infty} F_{\xi}(x_n) = F_{\xi}(x_0 + 0). \quad \blacktriangleright$$

8⁰. Імовірність того, що випадкова величина $\xi(\omega)$ набуває фіксоване значення x_0 , дорівнює стрибку функції розподілу в точці x_0 , тобто

$$P\{\omega: \xi(\omega) = x_0\} = F_{\xi}(x_0 + 0) - F_{\xi}(x_0).$$

◀ Оскільки

$$\{\omega: \xi(\omega) = x_0\} = \{\omega: \xi(\omega) \leq x_0\} \setminus \{\omega: \xi(\omega) < x_0\}$$

$$\text{і } \{\omega: \xi(\omega) < x_0\} \subset \{\omega: \xi(\omega) \leq x_0\},$$

то, за властивістю 3⁰ ймовірності,

$$P\{\omega: \xi(\omega) = x_0\} = P\{\omega: \xi(\omega) \leq x_0\} - P\{\omega: \xi(\omega) < x_0\} = \\ = F_{\xi}(x_0 + 0) - F_{\xi}(x_0). \quad \blacktriangleright$$

Якщо функція розподілу неперервна, то ймовірність випадкової величини набути будь-яке конкретне значення, дорівнює нулю.

Теорема (Ріса^{*}). Нехай деяка функція $F(x)$ визначена на всій числовій осі, не спадає, неперервна зліва і така, що $\lim_{x \rightarrow -\infty} F(x) = 0$ і $\lim_{x \rightarrow +\infty} F(x) = 1$. Тоді існує такий імовірнісний простір $(\Omega, \mathcal{F}, P)$ і така випадкова величина $\xi(\omega)$ на ньому, що функцією розподілу випадкової величини $\xi(\omega)$ буде функція $F(x)$, тобто $F_{\xi}(x) = F(x)$.

* Фрідьєши Ріс (Frigyés Riesz, 1880-1956) – угорський математик.

Фактично, ця теорема стверджує, що визначальними властивостями функції розподілу випадкової величини є монотонність, неперервність зліва, і нормування (рівність нулю на $-\infty$ і одиниці на $+\infty$).

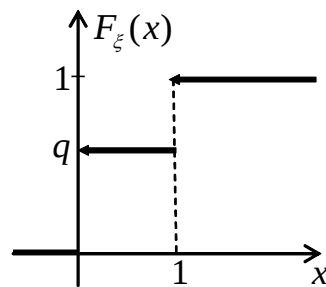
Приклад 1. Розглянемо випадкову величину $\xi(\omega)$, яка приймає два значення 0 і 1 з імовірностями q і p відповідно (таку випадкову величину називають індикатором події – кількість появ події в одному випробуванні). Це зручно записувати в таку таблицю:

$\xi(\omega)$	0	1
P	q	p

$$p > 0, \quad q > 0, \quad p + q = 1.$$

Тоді функція розподілу випадкової величини $\xi(\omega)$ і її графік мають такий вигляд:

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



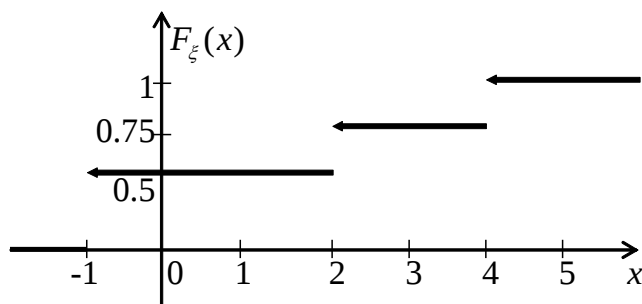
У точці $x=0$ стрибок значення функції величиною q – це і є $P\{\xi(\omega)=0\}$, а в точці $x=1$ стрибок величиною p і відповідно $P\{\xi(\omega)=1\}=p$.

Приклад 2. Нехай випадкова величина $\xi(\omega)$ набуває три значення -1, 2 і 4 з такими ймовірностями: $P\{\xi(\omega)=-1\}=0,5$, $P\{\xi(\omega)=2\}=0,25$ і $P\{\xi(\omega)=4\}=0,25$:

$\xi(\omega)$	-1	2	4
P	0,5	0,25	0,25

Функція розподілу випадкової величини $\xi(\omega)$ і її графік мають такий вигляд:

$$F_{\xi}(x) = \begin{cases} 0, & x \leq -1, \\ 0,5, & -1 < x \leq 2, \\ 0,75, & 2 < x \leq 4, \\ 1, & x > 4. \end{cases}$$



Приклад 3. Нехай випадкова величина $\xi(\omega)$ набуває n значень $x_1, x_2, \dots, x_n$ з імовірностями: $P\{\xi(\omega) = x_i\} = p_i, i = 1, 2, \dots, n$.

$\xi(\omega)$	x_1	x_2	x_3	$\dots$	x_n
P	p_1	p_2	p_3	$\dots$	p_n

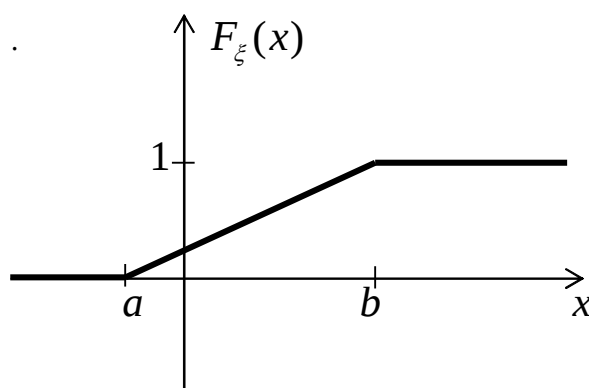
Тут $x_1 < x_2 < \dots < x_n$, $p_1 + p_2 + \dots + p_n = 1$. Запишемо $F_{\xi}(x)$ аналітично.

$$F_{\xi}(x) = \begin{cases} 0, & x \leq x_1, \\ \sum_{i: x_i < x} p_i, & x_1 < x \leq x_n, \\ 1, & x > x_n. \end{cases}$$

Приклад 4. Запишемо таку функцію $F_{\xi}(x)$ для $-\infty < a < b < +\infty$:

$$F_{\xi}(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & x > b. \end{cases}$$

Ця функція неперервна, не спадає, на $-\infty$ дорівнює нулю, на $+\infty$ дорівнює одиниці. Тому, за теоремою Ріса, це функція розподілу деякої випадкової величини.



З графіка видно, що випадкова величина набуває значення лише з відрізка $[a, b]$. Який би відрізок $[\alpha, \beta] \subset [a, b]$ не взяли, ймовірність випадкової величини набути значення з відрізка $[\alpha, \beta]$ дорівнює

$$P\{\omega: \alpha \leq \xi(\omega) < \beta\} = F_{\xi}(\beta) - F_{\xi}(\alpha) = \frac{\beta - a}{b - a} - \frac{\alpha - a}{b - a} = \frac{\beta - \alpha}{b - a}.$$

Але $P\{\omega: \xi(\omega) = \beta\} = 0$ (бо функція неперервна). Тому

$$P\{\omega: \alpha \leq \xi(\omega) \leq \beta\} = \frac{\beta - \alpha}{b - a} = \varphi(\beta - \alpha),$$

тобто ця ймовірність залежить лише від довжини відрізка. Отже, випадкова величина $\xi(\omega)$ розподілена рівномірно.

Якщо функція розподілу випадкової величини $\xi(\omega)$ має вигляд

$$F_{\xi}(x) = \begin{cases} 0, & x < a; \\ \frac{x - a}{b - a}, & a \leq x \leq b; \\ 1, & x > b. \end{cases}$$

то кажуть, що ця випадкова величина **розподілена рівномірно** на відріжку $[a, b]$.

Приклад 5. Нехай випадкова величина $\xi(\omega)$ розподілена рівномірно на $[-1, 1]$. Розглянемо випадкову величину $\eta = -\xi$. Знайдемо $F_{\eta}(x)$.

◀ З прикладу 4 відомо, що

$$F_{\xi}(x) = \begin{cases} 0, & x < -1, \\ \frac{x + 1}{2}, & -1 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

Тепер, на основі означення, запишемо $F_\eta(x)$:

$$F_\eta(x) = P\{\omega: \eta(\omega) < x\} = P\{\omega: -\xi(\omega) < x\} = P\{\omega: \xi(\omega) > -x\} = \\ = 1 - P\{\omega: \xi(\omega) \leq -x\} = 1 - P\{(\omega: \xi(\omega) < -x) \cup (\omega: \xi(\omega) = -x)\} =$$

події під знаком імовірності несумісні, тому:

$$= 1 - (P\{\omega: \xi(\omega) < -x\} + P\{\omega: \xi(\omega) = -x\}) =$$

але функція розподілу $\xi(\omega)$ неперервна, тому, за властивістю **8⁰**, останній доданок дорівнює нулю:

$$= 1 - P\{\omega: \xi(\omega) < -x\} = 1 - F_\xi(-x) =$$

$$= 1 - \begin{cases} 0, & -x < -1, \\ \frac{-x+1}{2}, & -1 \leq -x \leq 1, \\ 1, & -x > 1, \end{cases} = 1 - \begin{cases} 0, & x > 1, \\ \frac{-x+1}{2}, & -1 \leq x \leq 1, \\ 1, & x < -1, \end{cases} =$$

$$= \begin{cases} 1-0, & x > 1, \\ 1-\frac{-x+1}{2}, & -1 \leq x \leq 1, \\ 1-1, & x < -1, \end{cases} = \begin{cases} 1, & x > 1, \\ \frac{x+1}{2}, & -1 \leq x \leq 1, \\ 0, & x < -1, \end{cases} =$$

$$= \begin{cases} 0, & x < -1, \\ \frac{x+1}{2}, & -1 \leq x \leq 1, \\ 1, & x > 1, \end{cases} = \begin{cases} 0, & x < -1, \\ \frac{x+1}{2}, & -1 \leq x \leq 1, \\ 1, & x > 1, \end{cases} = F_\xi(x). \quad \blacktriangleright$$

Висновок: дві різні випадкові величини мають однакову функцію розподілу. Тобто, якщо функції розподілу двох випадкових величин однакові (співпадають), то випадкові величини можуть бути різні.

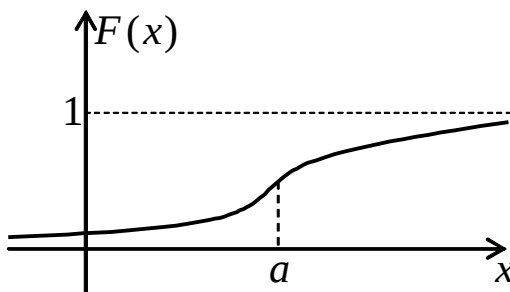
Приклад 6. Якщо функція розподілу випадкової величини $\xi(\omega)$ має вигляд

$$F_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{\frac{-(t-a)^2}{2\sigma^2}} dt,$$

то кажуть, що випадкова величина $\xi(\omega)$ **розподілена нормально** з параметрами a і σ ($a, \sigma \in \mathbb{R}, \sigma > 0$).

Те, що випадкова величина розподілена нормально з параметрами a і σ , записують так: $N(a, \sigma^2)$. Часто кажуть, що така випадкова величина **розподілена за законом Гаусса*** з параметрами a і σ .

Виникає питання: чи може функція $F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{\frac{-(t-a)^2}{2\sigma^2}} dt$ бути функцією розподілу? Графік цієї функції зображено на рисунку.



Вона неперервна, неспадна, і оскільки при $x = +\infty$ це функція Пуассона, то $F(x)$ є функцією розподілу деякої випадкової величини.

Для $N(0;1)$ таку функцію вже розглядали

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

і називали **функцією Лапласа**.

* Йоганн Карл Фрідріх Гаусс (Johann Carl Friedrich Gauß, 1777-1855) – німецький математик, астроном, геодезист та фізик.

Приклад 7. Нехай випадкова величина $\xi(\omega)$ має функцію розподілу $F_\xi(x)$. Розглянемо випадкову величину $\eta = a\xi + b$, $a, b \in \mathbb{R}$, $a \neq 0$. Якою є функція розподілу $F_\eta(x)$ випадкової величини η ?

◀ Можливі такі випадки:

1) $a > 0$. Тоді

$$F_\eta(x) = P\{\omega: \eta(\omega) < x\} = P\{\omega: a\xi + b < x\} = P\left\{\omega: \xi < \frac{x-b}{a}\right\} = F_\xi\left(\frac{x-b}{a}\right).$$

2) $a < 0$. Тоді

$$\begin{aligned} F_\eta(x) &= P\{\omega: a\xi + b < x\} = P\left\{\omega: \xi > \frac{x-b}{a}\right\} = \\ &= 1 - P\left\{\omega: \xi \leq \frac{x-b}{a}\right\} = 1 - F_\xi\left(\frac{x-b}{a} + 0\right). \end{aligned}$$

Як частковий випадок розглянемо такий приклад: $\xi(\omega)$ – розподілена рівномірно на відрізку $[-2, 3]$, тобто її функція розподілу має такий вигляд:

$$F_\xi(x) = \begin{cases} 0, & x < -2, \\ \frac{x+2}{5}, & -2 \leq x \leq 3, \\ 1, & x > 3. \end{cases}$$

Нехай $\eta_1 = 2\xi - 4$. Тоді η_1 розподілена рівномірно на відрізку $[2 \cdot (-2) - 4, 2 \cdot 3 - 4] = [-8, 2]$. Тому

$$F_{\eta_1}(x) = \begin{cases} 0, & x < -8, \\ \frac{x+8}{10}, & -8 \leq x \leq 2, \\ 1, & x > 2. \end{cases}$$

Якщо ж $\eta_2 = -3\xi + 5$, тоді η_2 – розподілена рівномірно на відрітку $[-4, 11]$. Отже,

$$F_{\eta_2}(x) = \begin{cases} 0, & x < -4, \\ \frac{x+4}{15}, & -4 \leq x \leq 11, \\ 1, & x > 11. \end{cases} \quad \blacktriangleright$$

Тобто при лінійних замінах ξ на η характер законів, переважно, не змінюється. Це правильно і для нормального закону.

Приклад 8. Нехай випадкова величина $\xi(\omega)$ має функцію розподілу $F_\xi(x)$. Розглядається випадкова величина $\eta = \xi^2$. Треба знайти функцію розподілу $F_\eta(x)$ випадкової величини η .

◀ Випадкова величина η набуває лише невід’ємні значення. Тому

$$F_\eta(x) = \begin{cases} 0, & x \leq 0 \\ ?, & x > 0 \end{cases}.$$

Розглянемо тепер $x > 0$. Тоді, за означенням,

$$\begin{aligned} F_\eta(x) &= P\{\eta < x\} = P\{\xi^2 < x\} = P\{-\sqrt{x} < \xi < \sqrt{x}\} = \\ &= P\left\{\left(-\sqrt{x} \leq \xi < \sqrt{x}\right) \setminus \left(-\sqrt{x} = \xi\right)\right\} = P\left\{-\sqrt{x} \leq \xi < \sqrt{x}\right\} - P\left\{\xi = -\sqrt{x}\right\} = \\ &= F_\xi(\sqrt{x}) - F_\xi(-\sqrt{x}) - F_\xi(-\sqrt{x} + 0) + F_\xi(-\sqrt{x}) = \\ &= F_\xi(\sqrt{x}) - F_\xi(-\sqrt{x} + 0). \end{aligned}$$

Тобто

$$F_\xi(x) = \begin{cases} 0, & x \leq 0, \\ F_\xi(\sqrt{x}) - F_\xi(-\sqrt{x} + 0), & x > 0. \end{cases} \quad \blacktriangleright$$

Приклад 9. Нехай випадкова величина ξ має функцію розподілу $F_\xi(x)$, причому $F_\xi(x)$ – неперервна, і монотонно зростає. Розглянемо випадкову величину $\eta = F_\xi(\xi)$. Треба знайти функцію розподілу $F_\eta(x)$ випадкової величини η .

◀ У загальному випадку при $0 \leq x \leq 1$ маємо

$$F_\eta(x) = P\{\eta < x\} = P\{F_\xi(\xi) < x\} = P\{\xi < F_\xi^{-1}(x)\} = F_\xi(F_\xi^{-1}(x)) = x.$$

Якщо ж $x < 0$, то

$$F_\eta(x) = P\{\eta < x\} = P\{F_\xi(\xi) < x\} = 0,$$

а при $x > 1$

$$F_\eta(x) = P\{\eta < x\} = P\{F_\xi(\xi) < x\} = 1.$$

Отже,

$$F_\eta(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

Але такий розподіл у рівномірно розподіленої на $[0, 1]$ випадкової величини.

Отже, η – це рівномірно розподілена на $[0, 1]$ випадкова величина. (Взагалі кажучи, досить вимагати лише неперервності $F_\xi(x)$ і не обов’язково монотонності).

Щоб зрозуміти, якою є випадкова величина η , розглянемо, наприклад, ξ з функцією розподілу

$$F_\xi(x) = \begin{cases} 0, & x \leq 0; \\ \frac{2}{\pi} \operatorname{arctg} x, & x > 0. \end{cases}$$

і тоді

$$\eta = \begin{cases} 0, & \xi \leq 0; \\ \frac{2}{\pi} \operatorname{arctg} \xi, & \xi > 0. \end{cases}$$

Функція розподілу $F_\eta(x)$ випадкової величини η при $\xi > 0$ і при $0 < x < 1$ має вигляд

$$\begin{aligned} F_\eta(x) &= P\{\eta < x\} = P\left\{\frac{2}{\pi} \operatorname{arctg} \xi < x\right\} = P\left\{\operatorname{arctg} \xi < \frac{\pi}{2} x\right\} = \\ &= P\left\{\xi < \operatorname{tg} \frac{\pi x}{2}\right\} = F_\xi\left(\operatorname{tg} \frac{\pi x}{2}\right) = \frac{2}{\pi} \operatorname{arctg}\left(\operatorname{tg} \frac{\pi x}{2}\right) = \frac{2}{\pi} \frac{\pi x}{2} = x. \end{aligned}$$

При $x \leq 0$ функція розподілу $F_\eta(x) = 0$, а при $x \geq 1$ $F_\eta(x) = 1$. ►

Приклад 10. Випадкова величина $\xi(\omega)$ розподілена нормально з параметрами $N(1, 4)$. Знайти функцію розподілу випадкової величини

$$\eta = \int_{-\infty}^{\xi} e^{-\frac{(x-1)^2}{8}} dx.$$

◀ Запишемо η у вигляді

$$\eta = \underbrace{\left(\frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{(x-1)^2}{8}} dx \right)}_{\eta_1} 2\sqrt{2\pi} = F_\xi(\xi) 2\sqrt{2\pi} = \eta_1 2\sqrt{2\pi},$$

але η_1 – рівномірно розподілена на $[0, 1]$ (це впливає з попереднього прикладу). Тому η – рівномірно розподілена на $[0, 2\sqrt{2\pi}]$, тобто

$$F_{\eta}(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{2\sqrt{2\pi}}, & 0 \leq x \leq 2\sqrt{2\pi}, \\ 1, & x > 2\sqrt{2\pi}. \end{cases} \quad \blacktriangleright$$

Приклад 11. Нехай випадкова величина $\xi(\omega)$ рівномірно розподілена на відрізку $[0, 1]$. Знайти функцію розподілу випадкової величини $\eta = \ln \frac{1}{\xi}$.

◀ Оскільки $\ln \frac{1}{\xi} > 0$, якщо $\xi \in (0, 1)$, то $P\{\eta < x\} = 0$ при $x \leq 0$.

Нехай тепер $x > 0$. Тоді

$$\begin{aligned} F_{\eta}(x) &= P\left\{\ln \frac{1}{\xi} < x\right\} = P\left\{\frac{1}{\xi} < e^x\right\} = P\{\xi > e^{-x}\} = \\ &= 1 - P\{\xi \leq e^{-x}\} = 1 - P\{\xi < e^{-x}\} = 1 - F_{\xi}(e^{-x}) = 1 - e^{-x}. \end{aligned}$$

Отже,

$$F_{\eta}(x) = 1 - e^{-x}.$$

Це так званий **показниковий розподіл** із параметром $\lambda = 1$ (про нього в §3 п.3.2). ▶

§3. Класифікація випадкових величин

Нехай задана деяка випадкова величина $\xi(\omega): \Omega \rightarrow \mathbb{R}$ на ймовірнісному просторі $(\Omega, \mathcal{F}, P)$.

Якщо випадкова величина $\xi(\omega)$ набуває не більше ніж зліченну кількість значень, то така випадкова величина називається **дискретною випадковою величиною**.

Ці значення, які вона набуває, позначаються $x_1, x_2, \dots, x_n, \dots$. Повинні бути задані й відповідні ймовірності $p_1, p_2, \dots, p_n, \dots$, де

$$p_i = P\{\xi(\omega) = x_i\}, \quad i = 1, 2, 3, \dots, \quad \sum_i p_i = 1.$$

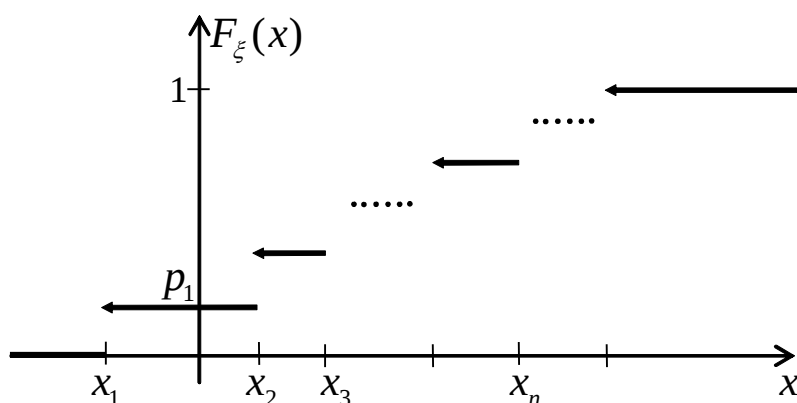
Тобто фактично задана така таблиця:

$\xi(\omega)$	x_1	x_2	$\dots$	x_n	$\dots$
P	p_1	p_2	$\dots$	p_n	$\dots$

Як бачимо, випадкова величина повністю задається і її функція розподілу може бути виписана так: $F_\xi(x) = \sum_{i: x_i \leq x} p_i$.

Отже, функція розподілу дискретної випадкової величини має не більше ніж зліченну кількість точок росту (x_0 називається **точкою росту** для $F_\xi(x)$, якщо для будь-якого $\varepsilon > 0$

$$F_\xi(x_0 + \varepsilon) - F_\xi(x_0 - \varepsilon) > 0).$$

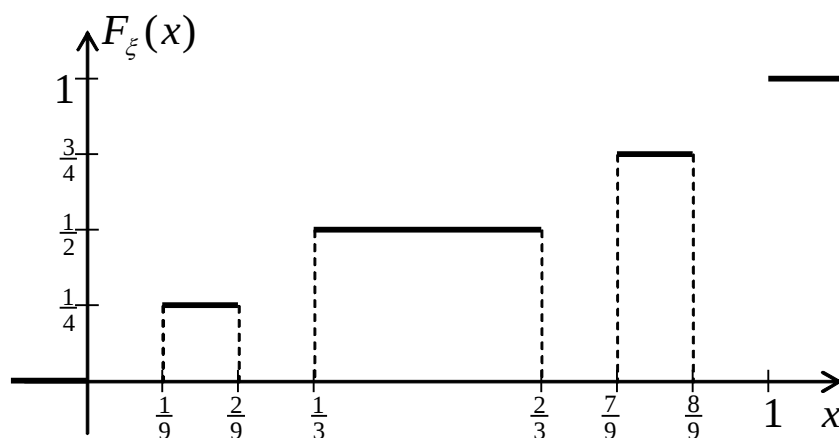


Вище вже розглядали функції, які мають континуальну множину точок росту (рівномірний розподіл, нормальний розподіл та інші). Але це були неперервні випадкові величини. Виникає питання: чи існують неперервні випадкові величини, які мають зліченну кількість точок росту? Відповідь: існують.

Приклад. Розглянемо випадкову величину ξ , яка набуває значення тільки між нулем і одиницею. Тому її функція розподілу задовольняє рівності $F_\xi(x) = 0$ при $x \leq 0$, $F_\xi(x) = 1$ при $x > 1$. На інтервалі $(0, 1)$ випадкова величина ξ набуває значення тільки у першій і останній його третинах, у кожній з імовірністю $\frac{1}{2}$. Отже, $F_\xi(x) = \frac{1}{2}$ при $\frac{1}{3} < x \leq \frac{2}{3}$. На інтервалах $\left(0, \frac{1}{3}\right)$ і $\left(\frac{2}{3}, 1\right)$ випадкова величина ξ знову може приймати значення тільки в першій і останній третинах інтервалів, у кожній з імовірністю $\frac{1}{4}$. Цим визначаються значення $F_\xi(x)$ ще на двох інтервалах:

$$F_\xi(x) = \frac{1}{4} \text{ при } \frac{1}{9} < x \leq \frac{2}{9},$$

$$F_\xi(x) = \frac{3}{4} \text{ при } \frac{7}{9} < x \leq \frac{8}{9}.$$



Далі на кожному з інтервалів, що залишилися, повторюється та

сама побудова, значення $F_{\xi}(x)$ визначається як константа, що дорівнює середньому арифметичному сусідніх, уже визначених значень $F_{\xi}(x)$, і цей процес триває до нескінченності. У точках, які не належать внутрішнім інтервалам, функція $F_{\xi}(x)$ довизначається за неперервністю. Випадкова величина ξ з функцією розподілу $F_{\xi}(x)$, визначеною описаним способом, не дискретна, оскільки її функція розподілу неперервна. Дійсно, сумарна довжина „внутрішніх” інтервалів, на яких $F_{\xi}(x)$ постійна, дорівнює

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots = \frac{1}{3} \sum_{k=0}^{\infty} \left(\frac{2}{3}\right)^k = \frac{1}{3} \frac{1}{1 - \frac{2}{3}} = 1,$$

так що точки росту функції $F_{\xi}(x)$ утворюють множину міри нуль (нульової міри Лебега^{*}).

Побудовану функцію називають **функцією Кантора**^{**}. Така функція має всі властивості функції розподілу, але у неї не більше ніж зліченна кількість точок росту. Такі функції називаються **сингулярними**.

Якщо функція розподілу випадкової величини абсолютно неперервна, то відповідна випадкова величина називається **абсолютно неперервною**, або просто **неперервною**.

(Детальніше про абсолютно неперервні функції див. Додаток А.)

У теорії ймовірностей вивчаються тільки неперервні і дискретні випадкові величини.

Абсолютно неперервну функцію розподілу можна представити у вигляді

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(u) du, \quad (*)$$

де $f_{\xi}(u) \geq 0$ – кусково-неперервна функція.

^{*} Анрі Леон Лебег (Henri Léon Lebesgue, 1875-1941) – французький математик.

^{**} Георг Кантор (Georg Cantor, 1845-1918) – німецький математик.

Тому іноді означення неперервної випадкової величини формують так: випадкова величина називається **неперервною**, якщо її функцію розподілу можна записати у вигляді (*).

Кусково-неперервна функція $f_{\xi}(x)$ називається **щільністю розподілу** випадкової величини.

Тому кажуть, що випадкова величина неперервна, якщо в неї існує щільність розподілу.

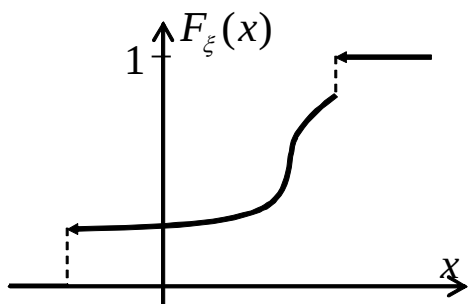
Теорема Лебега. Функцію розподілу будь-якої випадкової величини ξ можна в єдиний спосіб подати у вигляді суми

$$F_{\xi}(x) = F_{\partial\xi}(x) + F_{n\xi}(x) + F_{c\xi}(x),$$

трьох неспадних функцій $F_{\partial\xi}(x)$, $F_{n\xi}(x)$, $F_{c\xi}(x)$, де $F_{\partial\xi}(x)$ – дискретна складова, яка має вигляд $F_{\partial\xi}(x) = \sum_{i: x_i < x} p_i$, $\sum_i p_i \leq 1$, причому $p_i \geq 0$,

$F_{n\xi}(x)$ – неперервна складова, яка має вигляд $F_{n\xi}(x) = \int_{-\infty}^x f_{n\xi}(u) du$,

$\int_{-\infty}^{\infty} f_{n\xi}(u) du \leq 1$, причому $f_{n\xi}(x) \geq 0$, $F_{c\xi}(x)$ – сингулярна складова, що є неперервною функцією, множина точок росту якої має нульову міру Лебега.



У теорії ймовірностей вивчають такі випадкові величини, функції розподілу яких мають перші дві складові (дискретну й неперервну), тобто їх графіки такі, як на рисунку. Випадкові величини із сингулярними функціями розподілу не вивчають.

п. 3.1. Властивості щільності розподілу

1⁰. $f_{\xi}(x) \geq 0, \forall x \in \mathbb{R}$. Якщо б це було не так, то функція розподілу не була б монотонно неспадною.

2⁰. $\int_{-\infty}^{+\infty} f_{\xi}(u) du = 1.$

◀ Це повинно бути так, бо

$$\int_{-\infty}^{+\infty} f_{\xi}(u) du = \lim_{x \rightarrow \infty} \int_{-\infty}^x f_{\xi}(u) du = \lim_{x \rightarrow \infty} F_{\xi}(x) = 1$$

за властивістю **4⁰** функції розподілу. ▶

3⁰. $P\{a \leq \xi < b\} = \int_a^b f_{\xi}(u) du.$

◀ Ця властивість випливає з властивості **6⁰** функції розподілу:

$$P\{a \leq \xi < b\} = F_{\xi}(b) - F_{\xi}(a) = \int_{-\infty}^b f_{\xi}(u) du - \int_{-\infty}^a f_{\xi}(u) du = \int_a^b f_{\xi}(u) du. \quad \blacktriangleright$$

4⁰. Якщо функція розподілу $F_{\xi}(x)$ диференційовна в точці x , то

$$F'_{\xi}(x) = f_{\xi}(x).$$

Властивості **1⁰ – 4⁰** щільності розподілу часто використовуються при розв'язуванні задач.

Якщо задана інформація про випадкову величину у вигляді функції розподілу або щільності розподілу неперервної випадкової величини або таблиці значень і відповідних імовірностей дискретної випадкової величини, то кажуть, що задано **закон розподілу випадкової величини**.

Розглянемо приклади щільності розподілів.

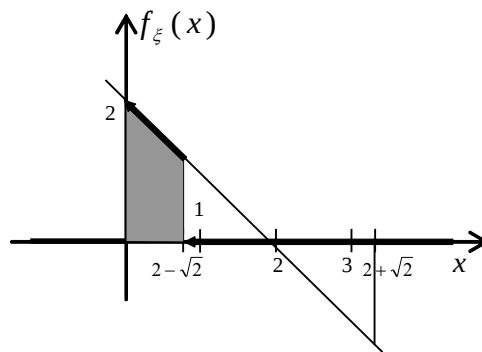
Приклад 1. Нехай $a > 0$ і

$$f_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ -x + 2, & 0 < x \leq a; \\ 0, & x > a. \end{cases}$$

Знайти константу a .

◀ Використаємо властивість 2⁰ щільності розподілу:

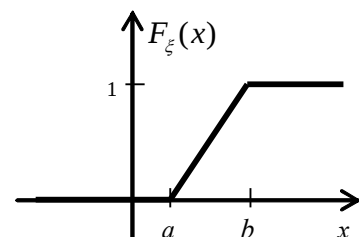
$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} f_{\xi}(u) du = \int_{-\infty}^0 f_{\xi}(u) du + \int_0^a f_{\xi}(u) du + \int_a^{+\infty} f_{\xi}(u) du = \\ &= \int_0^a (-u + 2) du = -\frac{a^2}{2} + 2a. \end{aligned}$$



Звідси $a^2 - 4a + 2 = 0$, $a_{1,2} = 2 \pm \sqrt{2}$. Яке ж вибрати a ? З властивості 1⁰ зрозуміло, що $a \neq 2 + \sqrt{2}$, бо $f_{\xi}(x)$ при $x = a = 2 + \sqrt{2}$ буде меншим нуля: $f_{\xi}(2 + \sqrt{2}) < 0$. Отже, $a = 2 - \sqrt{2}$. Площа заштрихованої частини дорівнює одиниці (за властивістю 2⁰). Відповідь: $a = 2 - \sqrt{2}$. ►

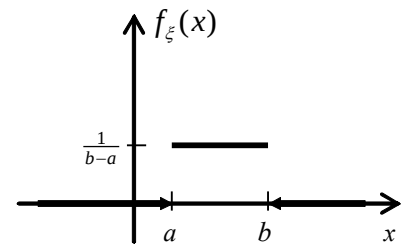
Приклад 2. Для рівномірного розподілу на $[a, b]$ була виписана функція розподілу $F_{\xi}(x)$

$$F_{\xi}(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & x > b. \end{cases}$$



Тоді одержуємо щільність розподілу

$$f_{\xi}(x) = \begin{cases} 0, & x < a; \\ \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & x > b. \end{cases}$$

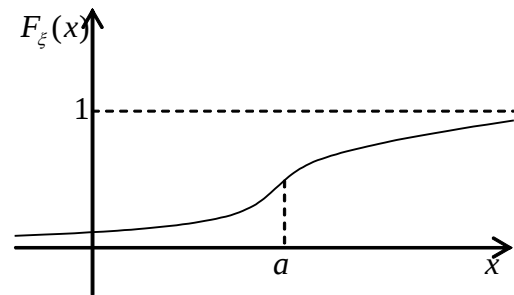


Можна $f_{\xi}(x)$ записати й інакше:

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

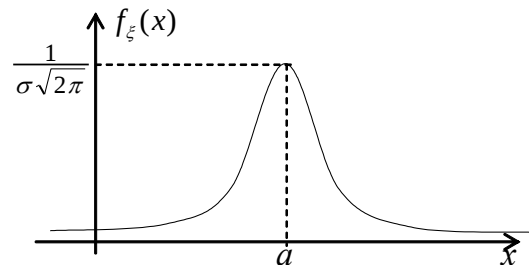
Приклад 3. Для **нормального розподілу** з параметрами a і σ функція розподілу має такий вигляд:

$$F_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-a)^2}{2\sigma^2}} du.$$



Тому для $N(a; \sigma^2)$ щільність розподілу записується так:

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



п. 3.2. Показниковий розподіл

Приклад 4. Нехай на проміжку часу $[0, \infty)$ працює деякий прилад, тобто розглядається прилад, який почав працювати у момент часу $t = 0$. Нехай він функціонує так, що ймовірність для приладу вийти з ладу на проміжку часу $(t, t + \Delta t)$, за умови, що цей прилад працює справно на проміжку $[0, t)$, має вигляд

$$P\{\text{вийти з ладу на } (t, t + \Delta t) \mid \text{працює справно на } [0, t)\} = \lambda \Delta t + o(\Delta t)$$

(λ називають доцільністю).

Позначимо через ξ таку випадкову величину – тривалість безвідмовної роботи приладу, тобто момент часу виходу з ладу приладу. Тоді, оскільки ξ – випадкова величина, то остання ймовірність така; $P\{t < \xi < t + \Delta t \mid \xi \geq t\} = \lambda \Delta t + o(\Delta t)$. Такі процеси зустрічаються, зокрема у телефонії. Знайдемо функцію розподілу випадкової величини ξ . Позначимо $Q(t) = P\{\xi \geq t\}$. Розглянемо

$$\begin{aligned} Q(t + \Delta t) &= P\{\xi \geq t + \Delta t\} = P\{(\xi \geq t + \Delta t) \cap (\xi \geq t)\} = \\ &= P\{\xi \geq t\} \cdot P\{\xi \geq t + \Delta t \mid \xi \geq t\} = \\ &= P\{\xi \geq t\} \cdot [1 - P\{t < \xi < t + \Delta t \mid \xi \geq t\}] = Q(t)[1 - \lambda \Delta t - o(\Delta t)]. \end{aligned}$$

Тобто

$$Q(t + \Delta t) - Q(t) = -\lambda Q(t) \Delta t - Q(t) \cdot o(\Delta t).$$

Поділимо останню рівність на Δt

$$\frac{Q(t + \Delta t) - Q(t)}{\Delta t} = -\lambda Q(t) - \frac{Q(t) \cdot o(\Delta t)}{\Delta t} = -\lambda Q(t) + \alpha(t)$$

і перейдемо до границі при $\Delta t \rightarrow 0$. Але $\alpha(t) \rightarrow 0$ при $\Delta t \rightarrow 0$, тому

при переході до границі одержуємо:

$$\frac{dQ(t)}{dt} = -\lambda Q(t)$$

– диференціальне рівняння з такою початковою умовою $Q(0) = 1$ (зрозуміло, що прилад колись вийде з ладу). Розв'язуючи цю задачу, одержуємо $Q(t) = e^{-\lambda t}$. Враховуючи позначення $Q(t) = P\{\xi \geq t\}$, маємо

$$P\{\xi < t\} = 1 - P\{\xi \geq t\} = 1 - Q(t) = \begin{cases} 0, & t \leq 0; \\ 1 - e^{-\lambda t}, & t > 0. \end{cases}$$

А ця імовірність є функцією розподілу випадкової величини ξ .

Означення. Якщо функція розподілу $F_\xi(x)$ випадкової величини ξ має вигляд

$$F_\xi(x) = \begin{cases} 0, & x \leq 0; \\ 1 - e^{-\lambda x}, & x > 0. \end{cases}$$

то кажуть, що випадкова величина ξ **розподілена за показниковим законом** (експоненціальним законом) із параметром λ .

Тоді можна записати щільність розподілу для випадкової величини, розподіленої за показниковим законом:

$$f_\xi(x) = \begin{cases} 0, & x \leq 0; \\ \lambda e^{-\lambda x}, & x > 0. \end{cases}$$

Виявляється, що випадкова величина, яка розподілена за показниковим законом, має властивість відсутності післядії (такою властивістю володіла дискретна випадкова величина, розподілена за геометричним законом). Ця властивість виражається так: нехай $0 < t < s$, тоді ймовірність

$$P\{t < \xi < s \mid \xi \geq t\} = P\{\xi < s - t\}.$$

Тобто: ймовірність того, що прилад вийде з ладу на відріжку $[t, s]$ (чи інтервалі (t, s)), якщо відомо, що до моменту t він працював справно, дорівнює ймовірності того, що прилад вийде з ладу на відріжку $[0, s - t]$.

◀ Перевіримо (доведемо) цю властивість:

$$P\{t < \xi < s \mid \xi \geq t\} = \frac{P\{(t < \xi < s) \cap (\xi \geq t)\}}{P(\xi \geq t)} = \frac{P(t < \xi < s)}{P(\xi \geq t)} =$$

функція $F_\xi(x)$ неперервна, тому

$$\begin{aligned} &= \frac{F_\xi(s) - F_\xi(t)}{1 - F_\xi(t)} = \frac{1 - e^{-\lambda s} - 1 + e^{-\lambda t}}{1 - 1 + e^{-\lambda t}} = \frac{e^{-\lambda t} (1 - e^{-\lambda(s-t)})}{e^{-\lambda t}} = \\ &= 1 - e^{-\lambda(s-t)} = P\{\xi < s - t\}. \quad \blacktriangleright \end{aligned}$$

Виявляється, що властивість відсутності післядії характерна для випадкової величини, розподіленої експоненціально, тобто це єдина неперервна випадкова величина, яка має таку властивість.

Доведення цього факту базується на використанні функціональних і різницевих рівнянь, зокрема використовується таке функціональне рівняння $Q(t + u) = Q(t) \cdot Q(u)$.

п. 3.3. Розподіл Коші*

Приклад 5. *Кажуть, що випадкова величина розподілена за законом Коші, якщо*

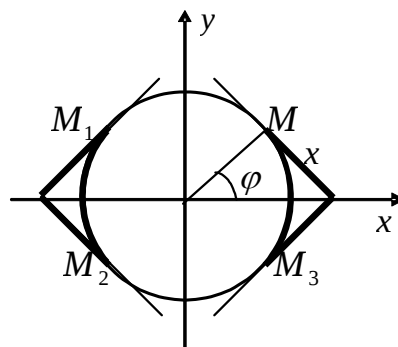
$$f_{\xi}(x) = \begin{cases} 0, & x < 0; \\ \frac{2}{\pi} \frac{1}{1+x^2}, & x \geq 0, \end{cases}$$

або

$$f_{\xi}(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \quad -\infty < x < \infty.$$

Звідки виникають такі випадкові величини? Розглянемо такий стохастичний експеримент. Нехай Ω – простір елементарних подій, точок одиничного кола, тобто $\Omega = \{(x, y): x^2 + y^2 = 1\}$. Ці точки рівноправні. Нехай задано число $x > 0$ і випадковим чином вибрана точка на колі. Яка ймовірність того, що довжина відрізка дотичної до кола у вибраній точці, від точки дотику до точки перетину з віссю Ox , не перевищує задане число x ?

Нехай випадкова величина ξ – довжина відрізка дотичної. Побудуємо точки M, M_1, M_2, M_3 на колі, довжини відрізків дотичних до кола у яких, від точки дотику до точки перетину з віссю Ox , дорівнюють x . Тоді подія $\xi < x$ відбудеться у тому і лише у тому випадку, коли точка буде вибрана на одній із двох дуг: M_1M_2 або M_3M .



* Огюстен Луї Коші (Augustin Louis Cauchy, 1789-1857) – французький математик.

Якщо простір елементарних подій звести до четвертини кола, то шукана ймовірність, з геометричних міркувань, дорівнює

$$P\{\xi < x\} = \frac{\varphi}{\frac{\pi}{2}} = \frac{\operatorname{arctg} x}{\frac{\pi}{2}} = \frac{2 \operatorname{arctg} x}{\pi}, \quad x > 0.$$

Але це функція розподілу, тому

$$F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{2}{\pi} \operatorname{arctg} x, & x > 0. \end{cases}$$

§4. Векторні випадкові величини

Нехай задано ймовірнісний простір $(\Omega, \mathcal{F}, P)$ і n випадкових величин на ньому $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$. Оскільки це випадкові величини на ймовірнісному просторі, то для будь-якого $i = 1, 2, \dots, n$ $\{\omega: \xi_i(\omega) < x\} \in \mathcal{F}$. Або навіть так: $\{\omega: \xi_i(\omega) \in B\} \in \mathcal{F}$, де B – довільна борелівська множина на $\mathcal{F}$. З випадкових величин $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$ можна скласти вектор $(\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega))$, що називається **випадковим вектором** – це вектор, компоненти якого є випадковими величинами.

Оскільки $\xi_i(\omega)$ – випадкові величини, то ω -множина

$$\{\omega: \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \dots, \xi_n(\omega) < x_n\},$$

де $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ ($x_1, x_2, \dots, x_n$ – дійсні числа), належить $\mathcal{F}$. Під цією ω -множиною розуміємо перетин таких множин:

$$\begin{aligned} & \{\omega: \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \dots, \xi_n(\omega) < x_n\} = \\ & = \{\omega: \xi_1(\omega) < x_1\} \cap \{\omega: \xi_2(\omega) < x_2\} \cap \dots \cap \{\omega: \xi_n(\omega) < x_n\}, \end{aligned}$$

що належать $\mathcal{F}$.

Це означає, що існує (визначена) ймовірність події

$$P\{\omega: \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \dots, \xi_n(\omega) < x_n\} = F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$$

і ця ймовірність називається **функцією сумісного розподілу** системи випадкових величин $\xi_1, \xi_2, \dots, \xi_n$.

Властивості функції сумісного розподілу

1⁰. $F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$ визначена для всіх $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

2⁰. $0 \leq F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) \leq 1$.

3⁰. $F_{\xi_1 \dots \xi_i \dots \xi_j \dots \xi_n}(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = F_{\xi_1 \dots \xi_j \dots \xi_i \dots \xi_n}(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$ – це впливає з означення відповідних ω -множин.

4⁰. $\lim_{x_1 \rightarrow +\infty} F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) = F_{\xi_2 \xi_3 \dots \xi_n}(x_2, x_3, \dots, x_n),$

$$\begin{aligned} \lim_{x_1 \rightarrow +\infty} F_{\xi_1 \xi_2 \dots \xi_{i-1} \xi_i \xi_{i+1} \dots \xi_n}(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) = \\ = F_{\xi_1 \xi_2 \dots \xi_{i-1} \xi_{i+1} \dots \xi_n}(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n). \end{aligned}$$

◀ Проведемо доведення властивості для двох змінних:

$$\begin{aligned} \lim_{x_1 \rightarrow +\infty} F_{\xi_1 \xi_2}(x_1, x_2) &= \lim_{x_1 \rightarrow +\infty} P\{\omega: (\xi_1(\omega) < x_1) \cap (\xi_2(\omega) < x_2)\} = \\ &= P\left\{\omega: \lim_{x_1 \rightarrow \infty} \{\xi_1(\omega) < x_1\} \cap \{\xi_2(\omega) < x_2\}\right\} = P\{\omega: \Omega \cap \{\xi_2(\omega) < x_2\}\} = \\ &= P\{\omega: \xi_2(\omega) < x_2\} = F_{\xi_2}(x_2). \quad \blacktriangleright \end{aligned}$$

Часто це записують так:

$$F_{\xi_1 \xi_2}(\infty, x_2) = F_{\xi_2}(x_2), \quad F_{\xi_1 \xi_2}(x_1, \infty) = F_{\xi_1}(x_1).$$

Зрозуміло, що $F_{\xi_1 \xi_2}(\infty, \infty) = 1$ (це можна довести, застосувавши почергово попередні рівності).

Для трьох змінних має місце, наприклад, така рівність:

$$F_{\xi_1 \xi_2 \xi_3}(x_1, \infty, x_3) = F_{\xi_1 \xi_3}(x_1, x_3).$$

Отже, якщо відома функція сумісного розподілу випадкових величин $\xi_1, \xi_2, \dots, \xi_n$, то легко знайти функції розподілу кожної з випадкових величин $\xi_1, \xi_2, \dots, \xi_n$ окремо. Наприклад, для ξ_1 можемо записати: $F_{\xi_1 \xi_2 \dots \xi_n}(x_1, \infty, \dots, \infty) = F_{\xi_1}(x_1)$. Постає питання про те, чи можна за відомими функціями розподілу $F_{\xi_1}(x_1), F_{\xi_2}(x_2), \dots, F_{\xi_n}(x_n)$ знайти функцію сумісного розподілу. Відповідь заперечна (пізніше у цьому переконаємося).

$$4'^0. \lim_{x_i \rightarrow -\infty} F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) = 0.$$

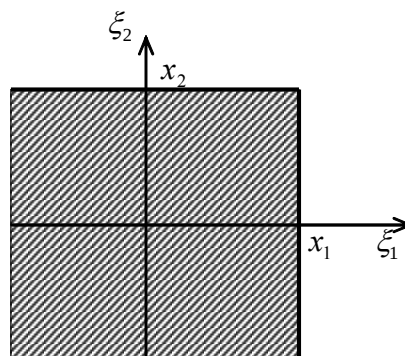
◀ Знову доведення проведемо для двох змінних:

$$\begin{aligned} \lim_{x_1 \rightarrow -\infty} F_{\xi_1 \xi_2}(x_1, x_2) &= \lim_{x_1 \rightarrow -\infty} P\{\omega: (\xi_1(\omega) < x_1) \cap (\xi_2(\omega) < x_2)\} = \\ &= P\left\{\omega: \left(\lim_{x_1 \rightarrow -\infty} (\xi_1(\omega) < x_1)\right) \cap (\xi_2(\omega) < x_2)\right\} = \\ &= P\{\omega: (\emptyset) \cap (\xi_2(\omega) < x_2)\} = P\{\omega: (\emptyset)\} = 0. \quad \blacktriangleright. \end{aligned}$$

5⁰. Функція сумісного розподілу неспадна по кожній зі змінних.

6⁰. Функція сумісного розподілу неперервна зліва по кожній зі змінних.

Подія, яка полягає у тому, що $\xi_1(\omega) < x_1$, а $\xi_2(\omega) < x_2$ (тобто $\{\omega: \xi_1(\omega) < x_1, \xi_2(\omega) < x_2\}$), може бути інтерпретована, як подія, яка полягає у тому, що випадковий вектор $(\xi_1(\omega), \xi_2(\omega))$ (тобто випадкова точка на площині) знаходиться у куті з вершиною у точці з



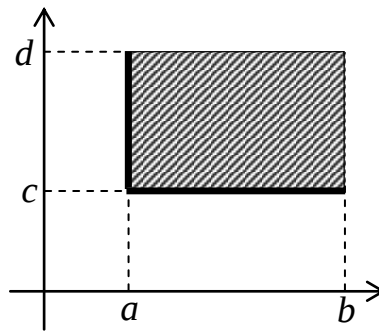
координатами (x_1, x_2) (тобто у від'ємному ортанті, зміщеному у точку з координатами (x_1, x_2)).

Тепер зрозуміло, що можна одержати аналог формули

$$P\{a \leq \xi(\omega) < b\} = F_{\xi}(b) - F_{\xi}(a).$$

7⁰. Імовірність

$$\begin{aligned} P\{a \leq \xi_1(\omega) < b, \quad c \leq \xi_2(\omega) < d\} &= \\ &= P\{\xi_1 < b, \quad \xi_2 < d\} - P\{\xi_1 < a, \quad \xi_2 < d\} - \\ &\quad - P\{\xi_1 < b, \quad \xi_2 < c\} + P\{\xi_1 < a, \quad \xi_2 < c\} = \\ &= F_{\xi_1 \xi_2}(b, \quad d) - F_{\xi_1 \xi_2}(a, \quad d) - F_{\xi_1 \xi_2}(b, \quad c) + F_{\xi_1 \xi_2}(a, \quad c). \end{aligned}$$



Пригадаємо: події незалежні, якщо ймовірність добутку подій дорівнює добутку їх імовірностей.

Означення 1. Дві випадкові величини $\xi_1(\omega)$, $\xi_2(\omega)$ називаються незалежними, якщо

$$P\{\omega: (\xi_1(\omega) < x_1) \cap (\xi_2(\omega) < x_2)\} = P\{\omega: \xi_1(\omega) < x_1\} \cdot P\{\omega: \xi_2(\omega) < x_2\}$$

для довільних $(x_1, x_2) \in \mathbb{R}^2$.

Тобто, якщо незалежні події записані у дужках.

Пригадаємо означення функції сумісного розподілу, й тоді останнє означення сформулюємо інакше.

Означення 1'. Дві випадкові величини $\xi_1(\omega)$, $\xi_2(\omega)$ називаються незалежними, якщо

$$F_{\xi_1 \xi_2}(x_1, x_2) = F_{\xi_1}(x_1) \cdot F_{\xi_2}(x_2),$$

тобто якщо функція сумісного розподілу дорівнює добутку функцій розподілу.

Означення 2. Випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ називаються попарно незалежними, якщо для будь-яких i, j ($i \neq j$) функція сумісного розподілу дорівнює добутку одновимірних функцій розподілу

$$F_{\xi_i \xi_j}(x_i, x_j) = F_{\xi_i}(x_i) \cdot F_{\xi_j}(x_j).$$

Означення 3. Випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ називаються незалежними в сукупності, якщо для будь-якого набору $(i_1, i_2, \dots, i_s)$, де $s \leq n$, такого, що $1 \leq i_1 < i_2 < \dots < i_s \leq n$, функція сумісного розподілу

$$F_{\xi_{i_1} \xi_{i_2} \dots \xi_{i_s}}(x_{i_1}, \dots, x_{i_s}) = \prod_{j=1}^s F_{\xi_{i_j}}(x_{i_j}).$$

Якщо випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ незалежні в сукупності, то вони і попарно незалежні. Якщо випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ попарно незалежні, то це ще не означає, що вони незалежні в сукупності (можна навести приклад Бернштейна для випадкових величин).

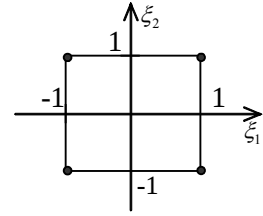
Якщо випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ незалежні в сукупності, то за їхніми одновимірними функціями розподілу можна побудувати функцію сумісного розподілу. У цьому випадку

$$F_{\xi_1 \dots \xi_n}(x_1, \dots, x_n) = \prod_{i=1}^n F_{\xi_i}(x_i).$$

Узагалі ж кажучи, за одновимірними функціями розподілу не можна побудувати функцію сумісного розподілу. Досить навести приклад, який показує, що це неможливо.

Приклад. Розглянемо систему випадкових величин (ξ_1, ξ_2) , таку, що випадковий вектор набуває чотири значення з однаковими ймовірностями:

(ξ_1, ξ_2)	(1, 1)	(1, -1)	(-1, -1)	(-1, 1)
P	0,25	0,25	0,25	0,25



Випишемо функцію сумісного розподілу:

$$F_{\xi_1 \xi_2}(x_1, x_2) = \begin{cases} 0, & x_1 \leq -1, \\ 0, & x_2 \leq -1, \\ \frac{1}{4}, & -1 < x_1 \leq 1, \quad -1 < x_2 \leq 1, \\ \frac{1}{2}, & -1 < x_1 \leq 1, \quad x_2 > 1, \\ \frac{1}{2}, & x_1 > 1, \quad -1 < x_2 \leq 1, \\ 1, & x_1 > 1, \quad x_2 > 1. \end{cases}$$

Розглянемо

$$\lim_{x_2 \rightarrow +\infty} F_{\xi_1 \xi_2}(x_1, x_2) = \begin{cases} 0, & x_1 \leq -1, \\ \frac{1}{2}, & -1 < x_1 \leq 1, \\ 1, & x_1 > 1 \end{cases} = F_{\xi_1}(x_1),$$

аналогічно,

$$F_{\xi_2}(x_2) = \lim_{x_1 \rightarrow +\infty} F_{\xi_1 \xi_2}(x_1, x_2) = \begin{cases} 0, & x_2 \leq -1, \\ \frac{1}{2}, & -1 < x_2 \leq 1, \\ 1, & x_2 > 1. \end{cases}$$

Вигляд функції $F_{\xi_1}(x_1)$ можна було б встановити й інакше, використовуючи закон розподілу випадкової величини ξ_1

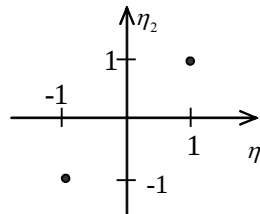
ξ_1	-1	1
P	0,5	0,5

Аналогічно, можна встановити і вигляд функції $F_{\xi_2}(x_2)$, використовуючи закон розподілу випадкової величини ξ_2

ξ_2	-1	1
P	0,5	0,5

Розглянемо ще одну систему випадкових величин (η_1, η_2) , яка набуває лише два значення з однаковими ймовірностями:

(η_1, η_2)	(-1, -1)	(1, 1)
P	0,5	0,5



Можна записати функцію сумісного розподілу $F_{\eta_1\eta_2}(x_1, x_2)$. Зрозуміло, що

$$F_{\eta_1\eta_2}(x_1, x_2) = \begin{cases} 0, & x_1 \leq -1, \\ 0, & x_2 \leq -1, \\ \frac{1}{2}, & -1 < x_1 \leq 1, \quad x_2 > -1, \\ \frac{1}{2}, & x_1 > 1, \quad -1 < x_2 \leq 1, \\ 1, & x_1 > 1, \quad x_2 > 1. \end{cases}$$

Отже, $F_{\eta_1\eta_2}(x_1, x_2) \neq F_{\xi_1\xi_2}(x_1, x_2)$. Знайдемо $F_{\eta_1}(x_1)$, $F_{\eta_2}(x_2)$:

η_1	-1	1
P	0,5	0,5

тому

$$F_{\eta_1}(x_1) = \begin{cases} 0, & x_1 \leq -1, \\ \frac{1}{2}, & -1 < x_1 \leq 1, \\ 1, & x_1 > 1 \end{cases} = F_{\xi_1}(x_1).$$

Аналогічно,

η_2	-1	1
P	0,5	0,5

$$F_{\eta_2}(x_2) = \begin{cases} 0, & x_2 \leq -1, \\ \frac{1}{2}, & -1 < x_2 \leq 1, \\ 1, & x_2 > 1 \end{cases} = F_{\xi_2}(x_2).$$

Отже,

$$F_{\xi_1}(x_1) = F_{\eta_1}(x_1), \quad F_{\xi_2}(x_2) = F_{\eta_2}(x_2), \quad F_{\xi_1\xi_2}(x_1, x_2) \neq F_{\eta_1\eta_2}(x_1, x_2).$$

А це і свідчить про те, що за одновимірними функціями розподілу побудувати функцію сумісного розподілу, взагалі кажучи, неможливо. Якби це можна було зробити, то за однаковими функціями розподілу одержали б однакові функції сумісного розподілу. Проте вони повинні бути різними.

§5. Щільність сумісного розподілу. Розподіл суми випадкових величин

Означення. Якщо функція сумісного розподілу $F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$ може бути зображена у вигляді

$$F_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f_{\xi_1 \xi_2 \dots \xi_n}(u_1, u_2, \dots, u_n) du_1 du_2 \dots du_n,$$

то $f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$ називається **щільністю сумісного розподілу**.

Властивості щільності сумісного розподілу

$$1^0. \quad f_{\xi_1 \xi_2 \dots \xi_n}(x_1, \dots, x_n) \geq 0.$$

$$2^0. \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\xi_1 \xi_2 \dots \xi_n}(u_1, u_2, \dots, u_n) du_1 du_2 \dots du_n = 1.$$

$$3^0. \quad P\{(\xi_1, \dots, \xi_n) \in B\} = \iiint_B f_{\xi_1 \dots \xi_n}(u_1, u_2, \dots, u_n) du_1 du_2 \dots du_n,$$

де B – борелівська множина.

4⁰. Якщо випадкові величини ξ_1 і ξ_2 незалежні (розглядається двовимірний випадок), то функція сумісного розподілу дорівнює добутку одновимірних функцій розподілу

$$F_{\xi_1 \xi_2}(x_1, x_2) = F_{\xi_1}(x_1) \cdot F_{\xi_2}(x_2).$$

Для абсолютно неперервних випадкових величин існують щільності розподілу $f_{\xi_1}(x_1)$, $f_{\xi_2}(x_2)$, тобто

$$F_{\xi_1}(x_1) = \int_{-\infty}^{x_1} f_{\xi_1}(u_1) du_1, \quad F_{\xi_2}(x_2) = \int_{-\infty}^{x_2} f_{\xi_2}(u_2) du_2.$$

Тоді функція сумісного розподілу

$$F_{\xi_1 \xi_2}(x_1, x_2) = \int_{-\infty}^{x_1} f_{\xi_1}(u_1) du_1 \cdot \int_{-\infty}^{x_2} f_{\xi_2}(u_2) du_2 = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{\xi_1}(u_1) f_{\xi_2}(u_2) du_1 du_2.$$

Отже, у цьому випадку виявляється, що

$$f_{\xi_1 \xi_2}(x_1, x_2) = f_{\xi_1}(x_1) \cdot f_{\xi_2}(x_2),$$

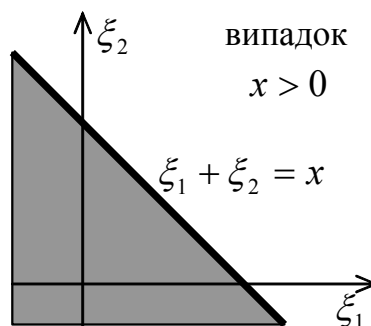
тобто: щільність сумісного розподілу двох незалежних абсолютно неперервних випадкових величин дорівнює добутку їх щільностей розподілу.

Приклад 1. Якщо система випадкових величин (випадковий вектор) розподілена рівномірно у крузі з межею $x^2 + y^2 = R^2$, то щільність повинна бути така: $f_{\xi_1 \xi_2}(x_1, x_2) = A$, де A – стала. Знайдемо A . За властивістю 2⁰, $\iint_{x^2+y^2 \leq R^2} A du_1 du_2 = 1$. Звідси $A\pi R^2 = 1$, тому

$$A = \frac{1}{\pi R^2}. \text{ Тобто}$$

$$f_{\xi_1 \xi_2}(x_1, x_2) = \begin{cases} 0, & x_1^2 + x_2^2 > R^2, \\ \frac{1}{\pi R^2}, & x_1^2 + x_2^2 \leq R^2. \end{cases}$$

Тепер розглянемо випадкову величину $\eta = \xi_1 + \xi_2 + \dots + \xi_n$, де $\xi_1, \xi_2, \dots, \xi_n$ – задані випадкові величини. Для спрощення розглянемо випадок $n = 2$. Тобто $\eta = \xi_1 + \xi_2$. Нехай задана щільність сумісного розподілу $f_{\xi_1 \xi_2}(x_1, x_2)$. Спробуємо знайти функцію розподілу $F_\eta(x)$ випадкової величини η .



$$F_{\eta}(x) = P\{\eta < x\} = P\{\xi_1 + \xi_2 < x\} = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{x-x_1} f_{\xi_1 \xi_2}(x_1, x_2) dx_2 =$$

проведемо заміну змінної $y = x_2 + x_1$

$$= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^x f_{\xi_1 \xi_2}(x_1, y - x_1) dy = \underbrace{\int_{-\infty}^x \int_{-\infty}^{\infty} f_{\xi_1 \xi_2}(x_1, y - x_1) dx_1 dy}_{\text{щільність розподілу } \eta \text{ (за означенням)}}$$

Тобто

$$f_{\eta}(y) = \int_{-\infty}^{\infty} f_{\xi_1 \xi_2}(x_1, y - x_1) dx_1$$

– щільність розподілу випадкової величини η і

$$F_{\eta}(x) = \int_{-\infty}^x f_{\eta}(y) dy.$$

Відмітимо, що у випадку, коли ξ_1 і ξ_2 незалежні, то, за властивістю 4⁰, щільність сумісного розподілу дорівнює добутку їх щільностей розподілу. Тому для випадкової величини η щільність розподілу може бути зображена у вигляді

$$f_{\eta}(y) = \int_{-\infty}^{\infty} f_{\xi_1}(x_1) \cdot f_{\xi_2}(y - x_1) dx_1 \quad \text{або} \quad f_{\eta}(y) = \int_{-\infty}^{\infty} f_{\xi_2}(x_2) \cdot f_{\xi_1}(y - x_2) dx_2.$$

Ці вирази називають **згорткою щільностей розподілу** і записують:

$$f_{\eta}(y) = f_{\xi_1} \otimes f_{\xi_2}.$$

Якщо випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ незалежні в сукупності, то

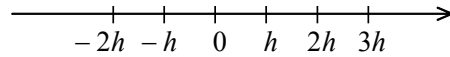
$$f_{\eta}(y) = \left(\left(f_{\xi_1} \otimes f_{\xi_2} \right) \otimes f_{\xi_3} \right) \otimes \dots \otimes f_{\xi_n}.$$

Отже, у випадку, коли випадкові величини незалежні в сукупності і кожна з них є абсолютно неперервною величиною (тобто існують щільності), то можна знайти щільність розподілу суми цих випадкових величин, яка дорівнює згортці їх щільностей розподілу.

Розглянемо більш загальний випадок: нехай задано дві незалежні випадкові величини ξ_1 і ξ_2 (наприклад, законом розподілу, функціями розподілу $F_{\xi_1}(x)$, $F_{\xi_2}(x)$), але не абсолютно неперервні, тобто щільностей не існує. Поставимо питання про знаходження функції розподілу $F_{\xi_1+\xi_2}(x)$ випадкової величини $\eta = \xi_1 + \xi_2$.

$$F_{\xi_1+\xi_2}(x) = F_{\eta}(x) = P\{\eta < x\} = P\{\xi_1 + \xi_2 < x\}.$$

Зафіксуємо деяке дійсне число $h > 0$ і всю числову вісь розіб'ємо на неперетинні відрізки довжиною h :



Простір елементарних подій Ω можна зобразити, наприклад, так:

$$\Omega = \bigcup_{k=-\infty}^{\infty} \{k \cdot h \leq \xi_1 + \xi_2 < (k+1)h\}.$$

Тоді, для будь-якого дійсного числа $h > 0$, ω -множину $\{\omega: \xi_1 + \xi_2 < x\}$ можна записати в такому вигляді:

$$\begin{aligned} \{\xi_1 + \xi_2 < x\} &= \{\xi_1 + \xi_2 < x\} \cap \Omega = \\ &= \{\xi_1 + \xi_2 < x\} \cap \bigcup_{k=-\infty}^{\infty} \{k \cdot h \leq \xi_1 + \xi_2 < (k+1)h\} = \\ &= \bigcup_{k=-\infty}^{\infty} (\{\xi_1 + \xi_2 < x\} \cap \{k \cdot h \leq \xi_1 + \xi_2 < (k+1)h\}) = \\ &= \bigcup_{k=-\infty}^{\infty} (\{\xi_1 < x - \xi_2\} \cap \{k \cdot h \leq \xi_2 < (k+1)h\}). \end{aligned}$$

Остання рівність одержана шляхом відщеплення змінних ξ_1 і ξ_2 . Цим самим події під знаком суми стали незалежними (одна ξ_1 , а друга ξ_2). Але, оскільки $\{\xi_1 + \xi_2 < x\} = \{\xi_1 < x - \xi_2\}$ і для всіх k виконуються нерівності $\{k \cdot h \leq \xi_2 < (k+1)h\}$, то

$$\{\xi_1 < x - (k+1)h\} \subset \{\xi_1 < x - \xi_2\} \quad \text{та} \quad \{\xi_1 < x - \xi_2\} \subset \{\xi_1 < x - kh\},$$

тобто

$$\{\xi_1 < x - (k+1)h\} \subset \{\xi_1 < x - \xi_2\} \subset \{\xi_1 < x - kh\},$$

а тому

$$\{\xi_1 < x - (k+1)h\} \subset \{\xi_1 + \xi_2 < x\} \subset \{\xi_1 < x - kh\}.$$

Отже, для всіх k

$$\begin{aligned} & \{\xi_1 < x - (k+1)h\} \cap \{kh \leq \xi_2 < (k+1)h\} \subset \\ & \subset \{\xi_1 + \xi_2 < x\} \cap \{kh \leq \xi_2 < (k+1)h\} \subset \\ & \subset \{\xi_1 < x - kh\} \cap \{kh \leq \xi_2 < (k+1)h\}. \end{aligned}$$

Просумуємо за всіма k , тоді

$$\begin{aligned} & \bigcup_{k=-\infty}^{+\infty} \{\xi_1 < x - (k+1)h\} \cap \{kh \leq \xi_2 < (k+1)h\} \subset \\ & \subset \bigcup_{k=-\infty}^{+\infty} \{\xi_1 + \xi_2 < x\} \cap \{kh \leq \xi_2 < (k+1)h\} \subset \\ & \subset \bigcup_{k=-\infty}^{+\infty} \{\xi_1 < x - kh\} \cap \{kh \leq \xi_2 < (k+1)h\}, \end{aligned}$$

або

$$\begin{aligned} & \bigcup_{k=-\infty}^{+\infty} \{\xi_1 < x - (k+1)h\} \cap \{kh \leq \xi_2 < (k+1)h\} \subset \{\xi_1 + \xi_2 < x\} \subset \\ & \subset \bigcup_{k=-\infty}^{+\infty} \{\xi_1 < x - kh\} \cap \{kh \leq \xi_2 < (k+1)h\}. \end{aligned}$$

Але множини $\{ih \leq \xi_2 < (i+1)h\}$ та $\{jh \leq \xi_2 < (j+1)h\}$ неперетинні при $i \neq j$, тобто відповідні події у ланцюжку включень несумісні, тому

під знаком суми доданки несумісні і ймовірність суми дорівнює сумі ймовірностей, тобто можна написати таке:

$$\begin{aligned} \sum_{k=-\infty}^{\infty} P\left(\{\xi_1 < x - (k+1)h\} \cap \{kh \leq \xi_2 < (k+1)h\}\right) &\leq P\{\xi_1 + \xi_2 < x\} \leq \\ &\leq \sum_{k=-\infty}^{\infty} P\left(\{\xi_1 < x - kh\} \cap \{kh \leq \xi_2 < (k+1)h\}\right). \end{aligned}$$

Крім цього, оскільки випадкові величини ξ_1 і ξ_2 незалежні, то зліва і справа під знаками суми маємо ймовірності добутоків незалежних випадкових величин. Тому останній ланцюжок нерівностей можна записати так:

$$\begin{aligned} \sum_{k=-\infty}^{\infty} P\{\xi_1 < x - (k+1)h\} \cdot P\{kh \leq \xi_2 < (k+1)h\} &\leq P\{\xi_1 + \xi_2 < x\} \leq \\ &\leq \sum_{k=-\infty}^{\infty} P\{\xi_1 < x - kh\} \cdot P\{kh \leq \xi_2 < (k+1)h\}. \end{aligned}$$

Звідси

$$\begin{aligned} \sum_{k=-\infty}^{\infty} F_{\xi_1}(x - (k+1)h) \cdot (F_{\xi_2}((k+1)h) - F_{\xi_2}(kh)) &\leq F_{\eta}(x) \leq \\ &\leq \sum_{k=-\infty}^{\infty} F_{\xi_1}(x - kh) \cdot (F_{\xi_2}((k+1)h) - F_{\xi_2}(kh)). \end{aligned}$$

Справа і зліва – інтегральні суми для інтеграла Стільтьєса^{*}.

Перейдемо до границі при $h \rightarrow 0$, тоді ліва і права суми збігаються до інтеграла

$$\int_{-\infty}^{\infty} F_{\xi_1}(x - t) dF_{\xi_2}(t),$$

^{*} *Томас Іоаннес Стільтьєс (Thomas Joannes Stieltjes, 1856-1894) – нідерландський математик.*

а тому

$$\int_{-\infty}^{\infty} F_{\xi_1}(x-t) dF_{\xi_2}(t) \leq F_{\eta}(x) \leq \int_{-\infty}^{\infty} F_{\xi_1}(x-t) dF_{\xi_2}(t).$$

Звідси

$$F_{\eta}(x) = \int_{-\infty}^{\infty} F_{\xi_1}(x-t) dF_{\xi_2}(t) = \int_{-\infty}^{\infty} F_{\xi_2}(x-t) dF_{\xi_1}(t).$$

Другу частину цієї рівності можна одержати, якщо перетворення (відщеплення змінних ξ_1 і ξ_2) проводити навпаки. Якщо цей запис розуміти як згортку, то можна записати $F_{\eta} = F_{\xi_1} \otimes F_{\xi_2}$. Тут $\otimes$ позначає, як і для щільностей, згортку функцій розподілу.

Для n незалежних у сукупності випадкових величин функція розподілу їх суми дорівнює згортці функцій розподілу випадкових величин:

$$F_{\xi_1 + \dots + \xi_n} = F_{\xi_1} \otimes F_{\xi_2} \otimes \dots \otimes F_{\xi_n}.$$

Припустимо тепер, що в однієї з випадкових величин (наприклад, у ξ_1) щільність розподілу існує. Тоді можна записати:

$$F_{\eta}(x) = \int_{-\infty}^{\infty} F_{\xi_2}(x-t) f_{\xi_1}(t) dt \quad (\text{це вже інтеграл Рімана}^*).$$

Можна показати, що у цьому випадку існує щільність розподілу випадкової величини η . Для цього досить провести лише заміну змінних (щоб x був верхньою межею інтегрування). Отже, для суми випадкових величин, щоб існувала щільність, досить, щоб існувала щільність для однієї з випадкових величин.

* *Георг Фрідріх Бернгард Ріман (Georg-Friedrich-Bernhard Riemann, 1826-1866) — німецький математик, механік і фізик.*

Приклад 2. Нехай задано дві незалежні дискретні випадкові величини ξ_1 і ξ_2 , які набувають цілочислові значення. Нехай

$$P\{\xi_1 = k\} = p_1(k), \quad k = 0, \pm 1, \pm 2, \dots,$$

$$P\{\xi_2 = k\} = p_2(k), \quad k = 0, \pm 1, \pm 2, \dots$$

Зрозуміло, що повинно бути

$$\sum_{k=-\infty}^{\infty} p_1(k) = 1, \quad \sum_{k=-\infty}^{\infty} p_2(k) = 1.$$

Розглянемо $\eta = \xi_1 + \xi_2$. Тоді

$$P\{\eta = k\} = \sum_{l+m=k} p_1(l)p_2(m) = \sum_{l=-\infty}^{\infty} p_1(l)p_2(k-l).$$

Приклад 3. Нехай задано дві незалежні дискретні випадкові величини ξ_1 і ξ_2 , які розподілені за законом Пуассона з параметрами λ_1 і λ_2 відповідно, тобто

$$p_1(k) = P\{\xi_1 = k\} = \frac{\lambda_1^k}{k!} e^{-\lambda_1}, \quad k = 0, 1, 2, \dots,$$

$$\sum_{k=0}^{+\infty} p_1(k) = 1,$$

$$p_2(k) = P\{\xi_2 = k\} = \frac{\lambda_2^k}{k!} e^{-\lambda_2}, \quad k = 0, 1, 2, \dots,$$

$$\sum_{k=0}^{+\infty} p_2(k) = 1.$$

Розглянемо суму $\xi_1 + \xi_2 = \eta$. Знайдемо (випишемо) закон розподілу для η , тобто $P\{\eta = k\}$, $k = 0, 1, 2, \dots$. Зрозуміло, що η набуває цілі значення.

$$\begin{aligned}
p(k) &= P\{\eta = k\} = P\{\xi_1 + \xi_2 = k\} = \sum_{l=-\infty}^{\infty} p_1(l) \cdot p_2(k-l) = \\
&= \sum_{l=-\infty}^{-1} \overset{=0}{p_1(l)} \cdot p_2(k-l) + \sum_{l=0}^{\infty} p_1(l) \cdot \overset{=0 \text{ при } k < l}{p_2(k-l)} = \\
&= \sum_{l=0}^k p_1(l) p_2(k-l) = \sum_{l=0}^k \frac{\lambda_1^l}{l!} e^{-\lambda_1} \cdot \frac{\lambda_2^{k-l}}{(k-l)!} e^{-\lambda_2} = \\
&= e^{-(\lambda_1 + \lambda_2)} \sum_{l=0}^k \frac{\lambda_1^l \cdot \lambda_2^{k-l}}{l!(k-l)!} = e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} \sum_{l=0}^k \frac{k!}{l!(k-l)!} \lambda_1^l \lambda_2^{k-l} = \\
&= e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} \sum_{l=0}^k C_k^l \lambda_1^l \lambda_2^{k-l} = e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} (\lambda_1 + \lambda_2)^k.
\end{aligned}$$

Тоді

$$P\{\eta = k\} = p(k) = \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)}.$$

Отже, сума двох незалежних випадкових величин ξ_1 і ξ_2 , які розподілені за законом Пуассона з параметрами λ_1 і λ_2 відповідно, розподілена за законом Пуассона з параметром λ , який дорівнює сумі параметрів випадкових величин $\lambda = \lambda_1 + \lambda_2$. Тобто закон Пуассона лінійний по параметру.

Приклад 3. Нехай дві випадкові величини ξ_1 і ξ_2 незалежні і розподілені рівномірно на відрізку $[a, b]$. Тобто

$$F_{\xi_1} = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b, \end{cases}$$

$$F_{\xi_2} = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

Функції розподілу однакові, але це не означає, що рівні випадкові величини: $\eta = \xi_1 + \xi_2 \neq 2\xi_1$. Знайдемо функцію розподілу (або щільність) випадкової величини $\eta = \xi_1 + \xi_2$. Шукаємо щільність. Було записано таке:

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1\xi_2}(x_1, x-x_1)dx_1, \quad f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1}(x_1) \cdot f_{\xi_2}(x-x_1)dx_1,$$

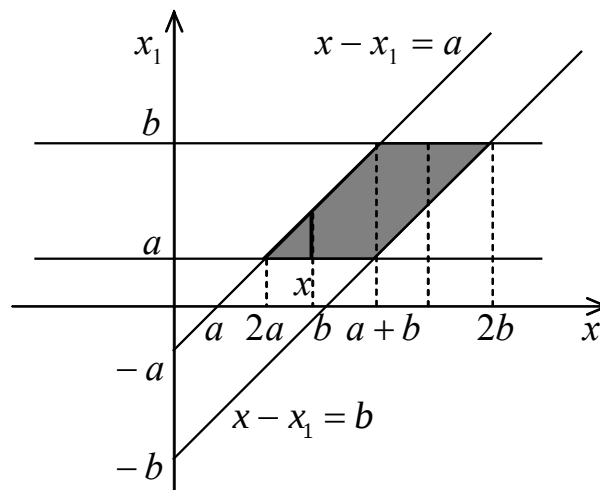
але у цьому випадку

$$f_{\xi_1}(x) = f_{\xi_2}(x) = \begin{cases} 0, & x \notin [a, b], \\ \frac{1}{b-a}, & x \in [a, b]. \end{cases}$$

Тому

$$f_{\xi_1+\xi_2}(x) = \frac{1}{b-a} \int_a^b f_{\xi_2}(x-x_1)dx_1,$$

межі інтегрування при різних значеннях x визначимо з рисунка:

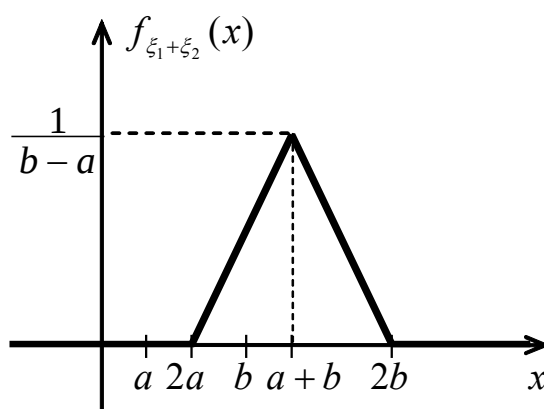


$$f_{\xi_1+\xi_2}(x) = \begin{cases} \frac{1}{b-a} \int_a^b 0 \cdot dx_1, & x < 2a, \\ \frac{1}{b-a} \int_a^{x-a} \frac{1}{b-a} dx_1, & 2a \leq x \leq a+b, \\ \frac{1}{b-a} \int_{x-b}^b \frac{1}{b-a} dx_1, & a+b < x \leq 2b, \\ \frac{1}{b-a} \int_a^b 0 \cdot dx_1, & x > 2b. \end{cases}$$

Тоді

$$f_{\xi_1+\xi_2}(x) = \begin{cases} 0, & x < 2a, \\ \frac{x-2a}{(b-a)^2}, & 2a \leq x \leq a+b, \\ \frac{2b-x}{(b-a)^2}, & a+b < x \leq 2b, \\ 0, & x > 2b. \end{cases}$$

Цей розподіл називається **розподілом Сімпсона*** (розподілом трикутника). Графік щільності розподілу η (розподілу Сімпсона) такий:



* Едвард Х'ю Сімпсон (Edward Hugh Simpson, 1922-2019) – британський статистик.

Функція розподілу випадкової величини $\eta = \xi_1 + \xi_2$ записується так:

$$F_{\xi_1 + \xi_2}(x) = \int_{-\infty}^x f_{\xi_1 + \xi_2}(u) du =$$

$$= \begin{cases} 0, & x < 2a, \\ \int_{2a}^x \frac{u - 2a}{(b - a)^2} du, & 2a \leq x \leq a + b, \\ \int_{2a}^{a+b} \frac{u - 2a}{(b - a)^2} du + \int_{a+b}^x \frac{2b - u}{(b - a)^2} du, & a + b < x \leq 2b, \\ 1, & x > 2b, \end{cases} =$$

$$= \begin{cases} 0, & x < 2a, \\ \frac{(x - 2a)^2}{2(b - a)^2}, & 2a \leq x \leq a + b, \\ \frac{2(a - b)^2 - (2b - x)^2}{2(b - a)^2}, & a + b < x \leq 2b, \\ 1, & x > 2b. \end{cases}$$

РОЗДІЛ 3

ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВИХ ВЕЛИЧИН

У попередньому розділі розглядалися кілька повних, вичерпних характеристик випадкових величин – так званих законів розподілу. Такими характеристиками були: для дискретних випадкових величин функція розподілу і ряд розподілу; для неперервних випадкових величин – функція розподілу і щільність розподілу.

Кожен закон розподілу є деякою функцією, і визначення такої функції повністю описує випадкову величину з імовірнісної точки зору.

Але у багатьох практичних питаннях немає необхідності характеризувати випадкову величину повністю, вичерпно. Часто буває достатньо вказати тільки окремі числові параметри, які певною мірою характеризують істотні властивості розподілу випадкової величини: наприклад, якесь середнє значення, навколо якого групуються можливі значення випадкової величини; якесь число, що характеризує міру розсіювання значень випадкової величини навколо середнього, і т.д. Такі характеристики, які дають змогу у стислій формі виразити найбільш істотні особливості розподілу, називаються числовими характеристиками випадкової величини.

У теорії ймовірностей і математичній статистиці використовуються багато різних числових характеристик, які мають різне призначення і різні сфери застосування. У цьому розділі розглянемо тільки деякі з них, які найчастіше використовуються.

§1. Числові характеристики випадкових величин

Нехай розглядається скалярна випадкова величина ξ на ймовірнісному просторі $(\Omega, \mathcal{F}, P)$ з функцією розподілу $F_\xi(x)$.

Означення 1. Математичним сподіванням випадкової величини ξ називається число $\mathbb{E}\xi$, яке дорівнює

$$\mathbb{E}\xi = \int_{-\infty}^{+\infty} x dF_\xi(x),$$

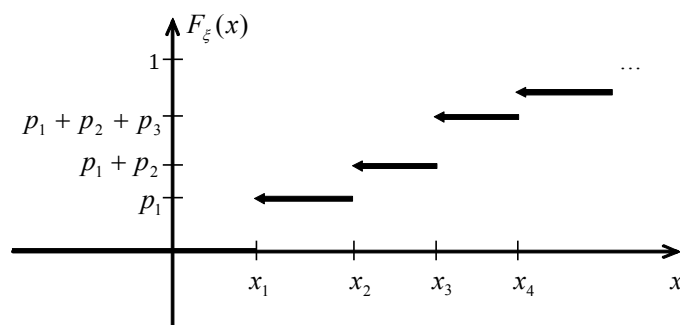
якщо такий інтеграл абсолютно збігається, тобто якщо існує інтеграл Стільтьєса*

$$\int_{-\infty}^{+\infty} |x| dF_\xi(x) < \infty.$$

Якщо ж інтеграл не збігається абсолютно, то кажуть, що математичного сподівання не існує.

Припустимо, що ξ — дискретна випадкова величина. Тоді функція розподілу є кусково-сталою.

ξ	x_1	x_2	x_3	$\dots$	x_k	$\dots$
P	p_1	p_1	p_3	$\dots$	p_k	$\dots$



* Означення і основні властивості інтеграла Стільтьєса див. у Додатку Б.

Якщо кусково-стала функція $F(x)$ змінює свої значення лише в точках $c_1, c_2, \dots, c_k, \dots$ і функція $f(x)$ неперервна на $[a, b]$, то інтеграл Стільтьєса $\int_{-\infty}^{+\infty} f(x) dF(x)$ – це така сума:

$$\sum_{k=1}^{\infty} f(c_k) [F(c_k + 0) - F(c_k - 0)].$$

Використавши у цій сумі $F(x) = F_{\xi}(x)$ і $f(x) = x$, $c_k = x_k$, одержуємо

$$\sum_{k=1}^{\infty} x_k [F_{\xi}(x_k + 0) - F_{\xi}(x_k - 0)].$$

Але стрибок функції розподілу в точці x_k , за властивістю **8⁰** функції розподілу, дорівнює

$$F_{\xi}(x_k + 0) - F_{\xi}(x_k - 0) = F_{\xi}(x_k + 0) - F_{\xi}(x_k) = P\{\xi = x_k\} = p_k.$$

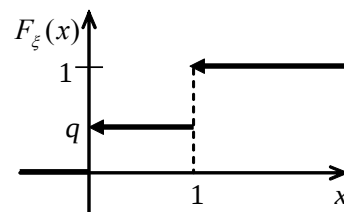
Отже, у випадку **дискретної випадкової величини математичне сподівання** дорівнює

$$\mathbb{E}\xi = \sum_k x_k p_k,$$

якщо цей ряд абсолютно збігається.

Приклад 1. Випадкова величина ξ набуває значення 0 і 1:

ξ	0	1
P	q	p



Таку випадкову величину називають ще індикатором події. Тоді математичне сподівання цієї випадкової величини дорівнює

$$\mathbb{E}\xi = 0 \cdot q + 1 \cdot p = p.$$

Приклад 2. Випадкова величина ξ набуває значення $0, 1, 2, \dots, n$ з імовірностями

$$P\{\xi = k\} = C_n^k p^k q^{n-k}, \quad p \geq 0, \quad q \geq 0, \quad p + q = 1$$

(розподіл Бернуллі, біноміальний розподіл).

$$\begin{aligned} \mathbb{E}\xi &= \sum_{k=0}^n x_k p_k = \sum_{k=0}^n k C_n^k p^k q^{n-k} = \sum_{k=1}^n k C_n^k p^k q^{n-k} = np \sum_{k=1}^n \frac{k}{n} C_n^k p^{k-1} q^{n-k} = \\ &= np \sum_{k=1}^n \frac{k}{n} \frac{n!}{k!(n-k)!} p^{k-1} q^{n-k} = np \sum_{k=1}^n \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} q^{(n-1)-(k-1)} = \\ &= np \sum_{m=0}^{n-1} \frac{(n-1)!}{m!(n-1-m)!} p^m q^{n-1-m} = np(p+q)^{n-1} = np. \end{aligned}$$

Отже, математичне сподівання випадкової величини ξ , розподіленої за законом Бернуллі, дорівнює

$$\mathbb{E}\xi = np.$$

Приклад 3. Нехай випадкова величина ξ розподілена за законом Пуассона з параметром λ , тобто

$$P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Тоді

$$\mathbb{E}\xi = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k \cdot k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} =$$

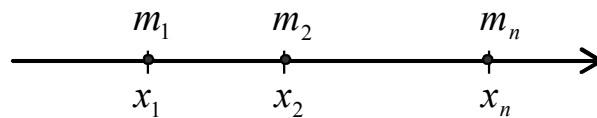
проведемо заміну $k-1 = m$, одержуємо ряд, який збігається до e^λ :

$$= \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda e^{-\lambda} e^\lambda = \lambda.$$

Для випадкової величини ξ , розподіленої за законом Пуассона з параметром λ , математичне сподівання дорівнює

$$\mathbb{E} \xi = \lambda.$$

Який зміст математичного сподівання для дискретної випадкової величини? Нехай в точках $x_1, x_2, \dots, x_n$ зосереджені маси $m_1, m_2, \dots, m_n$ відповідно.



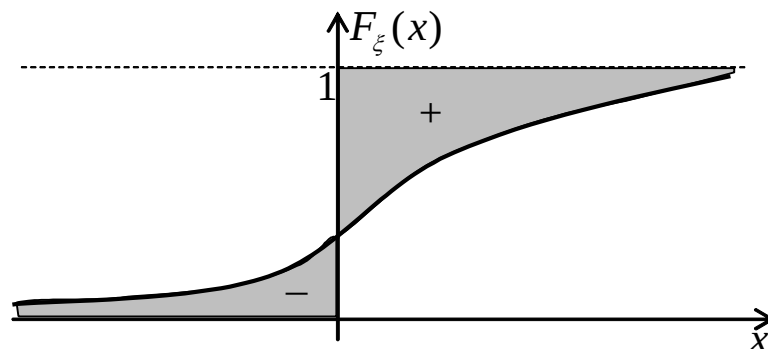
Центр ваги (мас) такої системи дорівнює

$$\frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{m_1 + m_2 + \dots + m_n} = x_1 \frac{m_1}{m} + x_2 \frac{m_2}{m} + \dots + x_n \frac{m_n}{m},$$

де $m = m_1 + m_2 + \dots + m_n$.

Якщо в точках числової осі зосереджені маси величиною p_i , то математичне сподівання є центром мас такої системи. Тобто математичне сподівання – це значення (якесь середнє), навколо якого групуються значення випадкової величини.

Використовуючи означення інтеграла Стільтьєса, можна дати просту геометричну інтерпретацію поняття математичного сподівання: математичне сподівання дорівнює різниці площ фігур, обмежених віссю ординат, прямою $y = 1$ і кривою $y = F_\xi(x)$ на інтервалі $(0, +\infty)$ і



між віссю абсцис, кривою $y = F_\xi(x)$ і віссю ординат на інтервалі

$(-\infty, 0)$. Зауважимо, що геометрична ілюстрація дозволяє записати математичне сподівання у такому вигляді:

$$\mathbb{E}\xi = -\int_{-\infty}^0 F_{\xi}(x)dx + \int_0^{+\infty} (1 - F_{\xi}(x))dx.$$

Якщо випадкова величина ξ неперервна зі щільністю розподілу $f_{\xi}(x)$, то інтеграл Стільтьєса в означенні 1 записується так:

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

Отже, для **випадкових величин абсолютно неперервного типу математичне сподівання дорівнює**

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

Це знову є центром мас, якщо розглядати стрижень з густиною $f_{\xi}(x)$.

Приклад 4. Нехай випадкова величина ξ розподілена рівномірно на відріжку $[a, b]$. Тобто

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

Тоді її математичне сподівання обчислюється так:

$$\mathbb{E}\xi = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{(b-a)2} = \frac{a+b}{2}.$$

Для випадкової величини, рівномірно розподіленої на відріжку $[a, b]$, математичне сподівання дорівнює

$$\mathbb{E}\xi = \frac{a+b}{2}.$$

Приклад 5. Нехай випадкова величина ξ розподілена нормально з параметрами $N(a; \sigma^2)$. Знайдемо $\mathbb{E}\xi$. Щільність розподілу $f_\xi(x)$ цієї абсолютно неперервної випадкової величини дорівнює

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

Тоді математичне сподівання $\mathbb{E}\xi$ обчислюється так:

$$\mathbb{E}\xi = \int_{-\infty}^{+\infty} x dF_\xi(x) = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx =$$

проведемо заміну $\frac{x-a}{\sigma} = t$, з якої одержуємо $x = \sigma t + a$, $dx = \sigma dt$, і тому

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

у першому з одержаних інтегралів підінтегральна функція непарна і тому він дорівнює нулю, а другий є інтегралом Пуассона і дорівнює $\sqrt{2\pi}$:

$$= \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \frac{a}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = a.$$

Отже, для нормально розподіленої випадкової величини ξ з параметрами $N(a; \sigma^2)$ математичне сподівання дорівнює

$$\mathbb{E}\xi = a.$$

Приклад 6. Нехай випадкова величина ξ розподілена за законом Коші. Щільність розподілу цієї абсолютно неперервної випадкової величини записується так:

$$f_\xi(x) = \frac{1}{\pi} \frac{1}{1+x^2} \quad -\infty < x < +\infty.$$

Перевіримо абсолютну збіжність інтеграла Стільтьєса з означення 1. Для цього розглянемо

$$\begin{aligned} \int_{-\infty}^{+\infty} |x| dF_{\xi}(x) &= \int_{-\infty}^{+\infty} |x| f_{\xi}(x) dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} |x| \frac{1}{1+x^2} dx = \frac{2}{\pi} \int_0^{+\infty} x \frac{1}{1+x^2} dx = \\ &= \frac{1}{\pi} \int_0^{+\infty} \frac{d(1+x^2)}{1+x^2} dx = \frac{1}{\pi} \ln(1+x^2) \Big|_0^{+\infty} = \frac{1}{\pi} \lim_{x \rightarrow +\infty} \ln(1+x^2) = +\infty. \end{aligned}$$

Отже, інтеграл Стільтьєса не є абсолютно збіжним. Це означає, що у випадкової величини ξ , розподіленої за законом Коші, не існує математичного сподівання.

Означення 2. *Дисперсією випадкової величини ξ називається число $D\xi$, яке дорівнює математичному сподіванню квадрата відхилення випадкової величини від її математичного сподівання:*

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

Дисперсія означає розсіювання і відіграє роль міри розсіювання значень випадкової величини навколо свого математичного сподівання.



Нехай розглядаються дві випадкові величини

ξ_1	-1	1
P	0,5	0,5

ξ_2	-5	5
P	0,5	0,5

Очевидно, що $\mathbb{E}\xi_1 = 0$ і $\mathbb{E}\xi_2 = 0$, але міра розсіювання (відхилення) цих випадкових величин навколо $\mathbb{E}\xi_1$ і $\mathbb{E}\xi_2$ різна.

Враховуючи означення математичного сподівання, можемо записати: $D\xi = \int_{-\infty}^{\infty} x dF_{\eta}(x)$, де $\eta = (\xi - \mathbb{E}\xi)^2$. В принципі $F_{\eta}(x)$ можна

знайти: $F_\eta(x) = P\{\eta < x\}$, але $F_\eta(x) = 0$ при $x \leq 0$, бо η набуває лише невід'ємні значення. Розглянемо тепер випадок $x > 0$:

$$\begin{aligned}
 F_\eta(x) &= P\{\eta < x\} = P\{(\xi - \mathbb{E}\xi)^2 < x\} = \\
 &= P\{-\sqrt{x} < \xi - \mathbb{E}\xi < \sqrt{x}\} = P\{-\sqrt{x} + \mathbb{E}\xi < \xi < \sqrt{x} + \mathbb{E}\xi\} = \\
 &= P\{-\sqrt{x} + \mathbb{E}\xi \leq \xi < \sqrt{x} + \mathbb{E}\xi\} - P\{\xi = -\sqrt{x} + \mathbb{E}\xi\} = \\
 &= F_\xi(\sqrt{x} + \mathbb{E}\xi) - F_\xi(-\sqrt{x} + \mathbb{E}\xi) - \\
 &\quad - F_\xi(-\sqrt{x} + \mathbb{E}\xi + 0) + F_\xi(-\sqrt{x} + \mathbb{E}\xi) = \\
 &= F_\xi(\sqrt{x} + \mathbb{E}\xi) - F_\xi(-\sqrt{x} + \mathbb{E}\xi + 0).
 \end{aligned}$$

Отже,

$$\begin{aligned}
 D\xi &= \int_{-\infty}^{\infty} x dF_\eta(x) = \int_0^{\infty} x dF_\eta(x) = \\
 &= \int_0^{\infty} x d \left[F_\xi(\sqrt{x} + \mathbb{E}\xi) - F_\xi(-\sqrt{x} + \mathbb{E}\xi + 0) \right] = \\
 &= \int_0^{\infty} x dF_\xi(\sqrt{x} + \mathbb{E}\xi) - \int_0^{\infty} x dF_\xi(-\sqrt{x} + \mathbb{E}\xi + 0) =
 \end{aligned}$$

проведемо такі заміни змінних: у першому інтегралі $\sqrt{x} + \mathbb{E}\xi = t$, тоді $x = (t - \mathbb{E}\xi)^2$, і в другому інтегралі $-\sqrt{x} + \mathbb{E}\xi + 0 = t$:

$$= \int_{\mathbb{E}\xi}^{\infty} (t - \mathbb{E}\xi)^2 dF_\xi(t) - \int_{\mathbb{E}\xi}^{-\infty} (t - \mathbb{E}\xi)^2 dF_\xi(t) = \int_{-\infty}^{\infty} (t - \mathbb{E}\xi)^2 dF_\xi(t).$$

Отже, одержали таку формулу для обчислення дисперсії:

$$D\xi = \int_{-\infty}^{\infty} (x - \mathbb{E}\xi)^2 dF_{\xi}(x).$$

Звідси:

$$\begin{aligned} D\xi &= \int_{-\infty}^{\infty} (x^2 - 2x\mathbb{E}\xi + (\mathbb{E}\xi)^2) dF_{\xi}(x) = \\ &= \int_{-\infty}^{\infty} x^2 dF_{\xi}(x) - 2\mathbb{E}\xi \int_{-\infty}^{\infty} x dF_{\xi}(x) + (\mathbb{E}\xi)^2 \int_{-\infty}^{\infty} dF_{\xi}(x). \end{aligned}$$

У другому доданку інтеграл дорівнює $\mathbb{E}\xi$, а у третьому – одиниці, бо є повною варіацією функції розподілу. Отже,

$$D\xi = \int_{-\infty}^{\infty} x^2 dF_{\xi}(x) - (\mathbb{E}\xi)^2.$$

Якщо випадкова величина дискретна, то одержується така формула для обчислення дисперсії випадкової величини ξ :

$$D\xi = \sum_k x_k^2 p_k - (\mathbb{E}\xi)^2 = \sum_k x_k^2 p_k - \left(\sum_k x_k p_k \right)^2.$$

Для випадкової величини абсолютно неперервного типу маємо:

$$D\xi = \int_{-\infty}^{\infty} x^2 f_{\xi}(x) dx - \left(\int_{-\infty}^{\infty} x f_{\xi}(x) dx \right)^2.$$

Означення 3. Моментом k -го порядку випадкової величини ξ відносно точки $x = a$ називається число

$$\nu_k(a) = \mathbb{E}(\xi - a)^k, \quad k = 1, 2, 3, \dots$$

Якщо $a = 0$, то $\nu_k(0) = \mathbb{E}(\xi)^k = \nu_k$ називається **початковим моментом** k -го порядку випадкової величини ξ .

Якщо $a = \mathbb{E}\xi$, то число $\nu_k(\mathbb{E}\xi) = \mathbb{E}(\xi - \mathbb{E}\xi)^k = \mu_k$ називається **центральним моментом** k -го порядку випадкової величини ξ .

Для центрального моменту k -го порядку можна записати таке:

$$\mu_k = \mathbb{E}(\xi - \mathbb{E}\xi)^k = \mathbb{E}\left(\sum_{i=0}^k C_k^i \xi^i (-1)^{k-i} (\mathbb{E}\xi)^{k-i}\right),$$

тобто центральні моменти виражаються через початкові.

Іноді ставиться питання: яким повинно бути число a , щоб $\nu_2(a)$ було мінімальним. Виявляється, що розв'язок цієї задачі $a = \mathbb{E}\xi$. Тому дисперсія $D\xi = \nu_2(\mathbb{E}\xi)$ є мінімальним значенням для моменту другого порядку $\nu_2(a)$.

У багатьох практичних задачах для обчислення математичного сподівання зручно використовувати наступну теорему.

Теорема. Нехай $\Xi = (\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega))$ — випадковий вектор із заданою щільністю сумісного розподілу $f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n)$ і випадкова величина $\eta = \varphi(\xi_1, \xi_2, \dots, \xi_n)$ є функцією від величин $\xi_1, \xi_2, \dots, \xi_n$. Тоді математичне сподівання випадкової величини η дорівнює

$$\mathbb{E}\eta = \int_{\mathbb{R}^n} \varphi(x_1, x_2, \dots, x_n) f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n,$$

якщо такий інтеграл абсолютно збігається, тобто якщо існує інтеграл

$$\int_{\mathbb{R}^n} |\varphi(x_1, x_2, \dots, x_n)| f_{\xi_1 \xi_2 \dots \xi_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

§2. Властивості числових характеристик

п. 2.1. Властивості математичного сподівання

1⁰. Якщо ймовірність $P\{\xi = C\} = 1$, то $E\xi = C$.

(Вужче ця властивість формулюється так: математичне сподівання сталої дорівнює сталій).

◀ Можна вважати, що ця випадкова величина дискретна з таким законом розподілу:

ξ	C	$*$
P	1	0

тут $*$ – довільне число і тоді $E\xi = C \cdot 1 + * \cdot 0 = C$. ▶

2⁰. Якщо ймовірність $P\{\xi \geq 0\} = 1$, то $E\xi \geq 0$.

(Якщо ξ набуває невід'ємні значення, то математичне сподівання невід'ємне).

◀ Для дискретних випадкових величин $\sum x_k p_k \geq 0$ при $x_k \geq 0$ і якщо ряд збігається.

Для абсолютно неперервних випадкових величин

$\int_{-\infty}^{\infty} x f_{\xi}(x) dx \geq 0$, бо $f_{\xi}(x) \geq 0$ при $x \geq 0$ і $f_{\xi}(x) = 0$ при $x < 0$. ▶

3⁰. Математичне сподівання суми випадкових величин дорівнює сумі математичних сподівань випадкових величин.

Для $\xi_1, \xi_2, E\xi_1, E\xi_2, \eta = \xi_1 + \xi_2$ маємо

$$E\eta = E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2.$$

◀ Проведемо доведення для дискретних випадкових величин.

Нехай випадкові величини ξ_1 і ξ_2 мають такі закони розподілу:

ξ_1	a_1	a_2	$\dots$	a_k	$\dots$
P	p_1	p_2	$\dots$	p_k	$\dots$

ξ_2	b_1	b_2	$\dots$	b_n	$\dots$
P	q_1	q_2	$\dots$	q_n	$\dots$

$$\sum_k p_k = 1,$$

$$\sum_n q_n = 1.$$

Для того, щоб знайти математичне сподівання суми, треба знайти, спочатку, розподіл суми $\xi_1 + \xi_2 = \eta$. Позначимо

$$p_{kn} = P\{\xi_1 = a_k, \xi_2 = b_n\} = P\{\eta = \xi_1 + \xi_2 = a_k + b_n\}$$

– закон розподілу η . Тоді, за означенням,

$$\begin{aligned} \mathbb{E} \eta &= \sum_{k,n=1}^{\infty} (a_k + b_n) p_{kn} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} (a_k + b_n) p_{kn} = \\ &= \sum_{k=1}^{\infty} a_k \underbrace{\sum_{n=1}^{\infty} p_{kn}}_{=p_k} + \sum_{n=1}^{\infty} b_n \underbrace{\sum_{k=1}^{\infty} p_{kn}}_{=q_n} = \\ &= \sum_{k=1}^{\infty} a_k p_k + \sum_{n=1}^{\infty} b_n q_n = \mathbb{E} \xi_1 + \mathbb{E} \xi_2. \quad \blacktriangleright \end{aligned}$$

Якщо існують хоча б два числа з чисел $\mathbb{E} \eta$, $\mathbb{E} \xi_1$, $\mathbb{E} \xi_2$, то існує й третє число.

Аналогічно отримується результат і для абсолютно неперервних випадкових величин.

У загальному випадку: математичне сподівання суми скінченної кількості випадкових величин дорівнює сумі їх математичних сподівань:

$$\mathbb{E} \left(\sum_{i=1}^n \xi_i \right) = \sum_{i=1}^n \mathbb{E} \xi_i.$$

4⁰. Якщо випадкові величини ξ_1 і ξ_2 **незалежні**, існують $\mathbb{E}\xi_1$ і $\mathbb{E}\xi_2$, то математичне сподівання добутку випадкових величин дорівнює добутку їх математичних сподівань:

$$\mathbb{E}(\xi_1 \cdot \xi_2) = \mathbb{E}\xi_1 \cdot \mathbb{E}\xi_2.$$

◀ Спочатку доведення проведемо для випадку абсолютно неперервних випадкових величин. Припустимо, що $f_{\xi_1 \xi_2}(x_1, x_2)$ – щільність сумісного розподілу випадкових величин ξ_1, ξ_2 . Тоді, враховуючи незалежність випадкових величин ξ_1 і ξ_2 , можемо записати таку рівність:

$$f_{\xi_1 \xi_2}(x_1, x_2) = f_{\xi_1}(x_1) f_{\xi_2}(x_2).$$

І тому, використовуючи результат сформульованої вище теореми, математичне сподівання $\mathbb{E}(\xi_1 \cdot \xi_2)$ дорівнює

$$\begin{aligned} \mathbb{E}(\xi_1 \cdot \xi_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{\xi_1}(x_1) f_{\xi_2}(x_2) dx_1 dx_2 = \\ &= \int_{-\infty}^{\infty} x_1 f_{\xi_1}(x_1) dx_1 \cdot \int_{-\infty}^{\infty} x_2 f_{\xi_2}(x_2) dx_2 = \mathbb{E}\xi_1 \cdot \mathbb{E}\xi_2. \end{aligned}$$

Тепер розглянемо випадок дискретних випадкових величин. Нехай випадкові величини ξ_1 і ξ_2 мають такі закони розподілу:

ξ_1	a_1	a_2	...	a_k	...
P	p_1	p_2	...	p_k	...

ξ_2	b_1	b_2	...	b_n	...
P	q_1	q_2	...	q_n	...

$$\sum_k p_k = 1,$$

$$\sum_n q_n = 1.$$

Тоді ймовірність того, що ξ_1 набуде значення a_k , а ξ_2 – значення b_n , дорівнює $p_k q_n$. За означенням математичного сподівання,

$$\begin{aligned}\mathbb{E}(\xi_1 \cdot \xi_2) &= \sum_{k,n=1}^{\infty} a_k b_n p_k q_n = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} a_k b_n p_k q_n = \\ &= \sum_{k=1}^{\infty} a_k p_k \cdot \sum_{n=1}^{\infty} b_n q_n = \mathbb{E}\xi_1 \cdot \mathbb{E}\xi_2. \quad \blacktriangleright\end{aligned}$$

Наступний приклад підтвердить важливість умови незалежності випадкових величин ξ_1 і ξ_2 :

Приклад. Нехай випадкова величина ξ_1 розподілена за законом

ξ_1	-2	2
P	0,5	0,5

а випадкова величина $\xi_2 = \xi_1^3$, тобто її закон розподілу такий:

ξ_2	-8	8
P	0,5	0,5

$\mathbb{E}\xi_1 = 0$, $\mathbb{E}\xi_2 = 0$, ξ_1 і ξ_2 – залежні. Запишемо закон розподілу добутку випадкових величин $\xi_1 \cdot \xi_2 = \xi_1^4$:

$\xi_1 \cdot \xi_2$	16
P	1

Тоді $\mathbb{E}(\xi_1 \cdot \xi_2) = \mathbb{E}(\xi_1^4) = 16 \neq \mathbb{E}\xi_1 \cdot \mathbb{E}\xi_2$.

5⁰. Постійний множник можна виносити за знак математичного сподівання:

$$\mathbb{E}(a\xi) = a \cdot \mathbb{E}\xi.$$

Це твердження очевидне, оскільки, якою б не була ξ , константу a і величину ξ можна розглядати як незалежні величини.

Використовуючи 5^0 , узагальнимо 3^0 :

$$3'^0. \mathbb{E}(a\xi_1 + b\xi_2) = a \cdot \mathbb{E}\xi_1 + b \cdot \mathbb{E}\xi_2.$$

6^0 . Якщо випадкові величини ξ_1 і ξ_2 такі, що $\xi_1 \leq \xi_2$, то $M\xi_1 \leq M\xi_2$.

◀ Дійсно, $\xi_2 - \xi_1$ – невід'ємна випадкова величина. Тому, згідно з властивістю 2^0 , $\mathbb{E}(\xi_2 - \xi_1) \geq 0$. Використовуючи властивість $3'^0$, одержуємо $\mathbb{E}\xi_2 \geq \mathbb{E}\xi_1$. ▶

7^0 . Легко зрозуміти, що $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$.

Це твердження випливає з нерівностей $-|\xi| \leq \xi \leq |\xi|$ і властивості 6^0 .

8^0 . Нехай випадкова величина $\xi \geq 0$, існує математичне сподівання $\mathbb{E}\xi$ і нехай $a > 0$ – деяке дійсне число. Тоді ймовірність

$$P\{\xi \geq a\} \leq \frac{\mathbb{E}\xi}{a}.$$

Ця нерівність називається **першою нерівністю Чебишова** або **нерівністю Маркова**.

◀ Розглянемо випадкову величину

$$\eta = \begin{cases} a, & \{\omega : \xi(\omega) \geq a\}, \\ 0, & \{\omega : \xi(\omega) < a\}. \end{cases}$$

Ясно, що $\xi \geq \eta$, звідси $\mathbb{E}\xi \geq \mathbb{E}\eta$ (якщо $\mathbb{E}\xi$ існує, то обов'язково існує $\mathbb{E}\eta$). Закон розподілу η такий:

$\eta:$	a	0
P	$P\{\xi \geq a\}$	$P\{\xi < a\}$

Тоді $\mathbb{E}\eta = a \cdot P\{\xi \geq a\}$, звідси $\mathbb{E}\xi \geq a \cdot P\{\xi \geq a\}$. ▶

9⁰. Нехай ξ – випадкова величина, $\eta = \varphi(\xi)$ – функція від ξ (наприклад, борелівська функція). Тоді математичне сподівання

$$\mathbb{E} \eta = \int_{-\infty}^{+\infty} \varphi(x) dF_{\xi}(x).$$

◀ За означенням, $\mathbb{E} \eta = \int_{-\infty}^{+\infty} x dF_{\eta}(x)$. Проводячи заміну змінних в інтегралі Стільтьєса, одержуємо **9⁰**.

Для дискретних випадкових величин треба записати закон розподілу η :

η	$\varphi(x_1)$	...
P	p_1	...

і, за означенням, знайти математичне сподівання випадкової величини η .

Для неперервних випадкових величин – провести заміну змінних в інтегралі Рімана. ►

При введенні в розгляд поняття дисперсії отримали таке співвідношення:

$$D\xi = \int_{-\infty}^{\infty} x^2 dF_{\xi}(x) - \left(\int_{-\infty}^{\infty} x dF_{\xi}(x) \right)^2$$

або у неперервному випадку:

$$\begin{aligned} D\xi &= \int_{-\infty}^{\infty} x^2 f_{\xi}(x) dx - \left(\int_{-\infty}^{\infty} x f_{\xi}(x) dx \right)^2 = \\ &= \int_{-\infty}^{\infty} x^2 f_{\xi}(x) dx - (\mathbb{E} \xi)^2 \stackrel{\text{за } 9^0}{=} \mathbb{E} \xi^2 - (\mathbb{E} \xi)^2. \end{aligned}$$

Отже, крім означення

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

отримали ще й формулу

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

Цю формулу можна одержати й інакше (використовуючи властивості математичного сподівання):

$$\begin{aligned} D\xi &= \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{E}(\xi^2 - 2\xi \cdot \mathbb{E}\xi + (\mathbb{E}\xi)^2) = \\ &= \mathbb{E}\xi^2 - 2\mathbb{E}\xi \cdot \mathbb{E}\xi + (\mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2. \end{aligned}$$

Крім цього, дисперсія – це мінімальне значення моменту другого порядку, тобто функції $\mathbb{E}(\xi - a)^2$. Дійсно,

$$\mathbb{E}(\xi - a)^2 = \mathbb{E}(\xi^2 - 2\xi a + a^2) = \mathbb{E}\xi^2 - 2a \cdot \mathbb{E}\xi + a^2 = I(a).$$

Яким треба вибрати a , щоб цей вираз (момент другого порядку) був мінімальним? Одержуємо: $I'(a) = -2 \cdot \mathbb{E}\xi + 2a$. З умови $I'(a) = 0$ випливає, що при $a = \mathbb{E}\xi$ досягається, як неважко переконатися, мінімальне значення. Тобто, як було сказано вище, мінімальний момент відносно значення a досягається при $a = \mathbb{E}\xi$. Тоді цей момент дорівнює $\mathbb{E}(\xi - \mathbb{E}\xi)^2$ і, за означенням, називається дисперсією.

п. 2.2. Властивості дисперсії

1⁰. Для будь-якої випадкової величини ξ дисперсія $D\xi \geq 0$. Причому $D\xi = 0$ тоді і тільки тоді, коли існує така стала C , що $P\{\xi = C\} = 1$ (тобто дисперсія сталої дорівнює нулю, можна довести і навпаки: якщо $D\xi = 0$, то ξ – стала з імовірністю одиниця).

2⁰. $D(a\xi + b) = a^2 \cdot D\xi$.

◀ За означенням дисперсії,

$$\begin{aligned} D(a\xi + b) &= \mathbb{E}(a\xi + b - \mathbb{E}(a\xi + b))^2 = \mathbb{E}(a\xi + b - a \cdot \mathbb{E}\xi - b)^2 = \\ &= \mathbb{E}(a(\xi - \mathbb{E}\xi))^2 = a^2 \cdot \mathbb{E}(\xi - \mathbb{E}\xi)^2 = a^2 \cdot D\xi. \quad \blacktriangleright \end{aligned}$$

Ця властивість вказує на те, що дисперсія $a\xi + b$ не залежить від b : бо випадкова величина ξ зміщується на b і одержується випадкова величина $\xi + b$, а дисперсія – це міра розсіювання відносно математичного сподівання (відповідного).



Означення 4. Число $\sigma_\xi^\xi = \sigma_\xi = \sqrt{D\xi}$ називається *середнім квадратичним відхиленням* випадкової величини ξ .

Тоді можна записати властивість, аналогічну до 2⁰.

2'⁰. $\sigma_{a\xi+b} = |a| \sigma_\xi$.

3⁰. Дисперсія суми або різниці незалежних випадкових величин дорівнює сумі дисперсій випадкових величин: якщо ξ , η – незалежні, тоді

$$D(\xi \pm \eta) = D\xi + D\eta.$$

◀ За означенням дисперсії,

$$\begin{aligned} D(\xi \pm \eta) &= \mathbb{E}(\xi \pm \eta - \mathbb{E}(\xi \pm \eta))^2 = \mathbb{E}((\xi - \mathbb{E}\xi) \pm (\eta - \mathbb{E}\eta))^2 = \\ &= \mathbb{E}(\xi - \mathbb{E}\xi)^2 + \mathbb{E}(\eta - \mathbb{E}\eta)^2 \pm 2\mathbb{E}(\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta) = \end{aligned}$$

якщо ξ, η – незалежні, то $\xi - \mathbb{E}\xi$ та $\eta - \mathbb{E}\eta$ – незалежні і тоді, за властивістю **4⁰** математичного сподівання, одержуємо:

$$= D\xi + D\eta \pm 2 \cdot \underbrace{\mathbb{E}(\xi - \mathbb{E}\xi)}_{=0} \cdot \underbrace{\mathbb{E}(\eta - \mathbb{E}\eta)}_{=0} = D\xi + D\eta. \quad \blacktriangleright$$

Дисперсія суми **попарно незалежних** випадкових величин (досить саме попарної незалежності і не обов'язково незалежності у сукупності) дорівнює сумі дисперсій випадкових величин.

Сформульовані властивості **1⁰–3⁰** є важливим інструментом обчислення дисперсій випадкових величин.

Приклад 1. Нехай випадкова величина розподілена за біноміальним законом (законом Бернуллі), тобто

$$P\{\xi = k\} = C_n^k p^k q^{n-k}, \quad k = \overline{0, n} \quad \text{і} \quad p + q = 1, \quad p > 0, \quad q > 0.$$

Було вже обчислено $M\xi = np$. Обчислимо $D\xi$.

$$D\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{E}(\xi - np)^2 = \sum_{k=0}^n (k - np)^2 C_n^k p^k q^{n-k}.$$

Можна тут продовжити обчислення, але можна скористатися робочою формулою обчислення дисперсії

$$D\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \sum_{k=0}^n k^2 C_n^k p^k q^{n-k} - (np)^2$$

і тут продовжити обчислення. Але виберемо такий шлях: зіставимо випадкову величину ξ з послідовністю незалежних випробовувань.

Нехай таких випробувань n . Введемо в розгляд випадкові величини ξ_i :

$$\xi_i = \begin{cases} 1, & \text{якщо подія } A \text{ відбулася в } i\text{-му експерименті,} \\ 0, & \text{якщо подія } A \text{ не відбулася в } i\text{-му експерименті,} \end{cases} \quad i=1,2,\dots,n.$$

Випадкові величини- ξ_i – незалежні, однаково розподілені:

ξ_i	0	1
P	q	p

Тоді для $i=1,2,\dots,n$

$$\mathbb{E}\xi_i = 0 \cdot q + 1 \cdot p = p,$$

$$D\xi_i = \mathbb{E}\xi_i^2 - (\mathbb{E}\xi_i)^2 = 0^2 \cdot q + 1^2 \cdot p - p^2 = p - p^2 = p(1-p) = p \cdot q.$$

Легко зрозуміти, що випадкова величина $\xi = \xi_1 + \xi_2 + \dots + \xi_n$. І тепер застосовуємо властивості математичного сподівання і дисперсії для суми випадкових величин:

$$\mathbb{E}\xi = \mathbb{E}\xi_1 + \mathbb{E}\xi_2 + \dots + \mathbb{E}\xi_n = n \mathbb{E}\xi_i = np$$

і

$$D\xi = D\xi_1 + D\xi_2 + \dots + D\xi_n = nD\xi_i = npq.$$

Отже, для біноміального закону розподілу (закону Бернуллі)

$$D\xi = npq, \quad \sigma_\xi = \sqrt{npq}.$$

Приклад 2. Нехай випадкова величина ξ розподілена рівномірно на $[a, b]$. Це означає, що

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

Крім цього, знаємо, що $\mathbb{E}\xi = \frac{a+b}{2}$. Знайдемо $D\xi$.

$$\begin{aligned} D\xi &= \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \int_a^b x^2 \frac{1}{b-a} dx - \left(\frac{a+b}{2}\right)^2 = \\ &= \frac{1}{b-a} \frac{b^3 - a^3}{3} - \frac{(a+b)^2}{4} = \frac{b^2 + ab + a^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \\ &= \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12}. \end{aligned}$$

Отже, при $b > a$

$$D\xi = \frac{(b-a)^2}{12}, \quad \sigma_\xi = \frac{b-a}{2\sqrt{3}}$$

для рівномірно розподіленої на $[a, b]$ випадкової величини ξ .

Рівномірний закон розподілу визначається числами a і b . З іншого боку, знаючи $\mathbb{E}\xi$ і $D\xi$ ці числа (a і b) можна знайти. Отже, рівномірний закон розподілу повністю визначається двома моментами (першим початковим і другим центральним). Якщо у загальному випадкові величини визначаються функцією або щільністю, то для рівномірного закону розподілу випадкова величина визначається двома моментами.

Приклад 3. Нехай випадкова величина ξ нормально розподілена з параметрами $N(a; \sigma^2)$, $\sigma > 0$. Було знайдено $\mathbb{E}\xi = a$. Обчислимо $D\xi$.

$$D\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \int_{-\infty}^{\infty} x^2 f_\xi(x) dx - a^2 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx - a^2 =$$

проведемо заміну змінних $\frac{x-a}{\sigma} = t$, тоді $x = \sigma t + a$, $dx = \sigma \cdot dt$ і одержуємо:

$$\begin{aligned}
&= \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma^2 t^2 + 2a\sigma t + a^2) e^{-\frac{t^2}{2}} dt - a^2 = \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt + \underbrace{\frac{2a\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} dt}_{=0, \text{ бо підінтегр. функція непарна}} + \underbrace{\frac{a^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt}_{=\sqrt{2\pi}} - a^2 = \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt = \frac{\sigma^2}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} t \cdot \underbrace{e^{-\frac{t^2}{2}} dt^2}_{dv} = \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right) = \frac{\sigma^2}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) = \sigma^2.
\end{aligned}$$

Отже, для нормально розподіленої випадкової величини з параметрами $N(a; \sigma^2)$

$$D\xi = \sigma^2, \quad \sigma_{\xi} = \sigma.$$

Тобто нормальний закон розподілу, як і рівномірний, повністю визначається двома параметрами – математичним сподіванням і дисперсією. Існують такі випадкові величини, що визначаються лише одним параметром (наприклад, закон Пуассона визначається математичним сподіванням).

4⁰. Для будь-якого $\varepsilon > 0$ імовірність того, що випадкова величина ξ відрізняється від свого математичного сподівання не менше ніж на ε , не перевищує дисперсії випадкової величини ξ , поділеної на ε^2 :

$$P\{|\xi - \mathbb{E}\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}.$$

Ця нерівність називається **нерівністю Чебишова** (другою нерівністю Чебишова).

◀ Розглянемо випадкову величину $\eta = (\xi - \mathbb{E}\xi)^2$. Зрозуміло, що $\eta = (\xi - \mathbb{E}\xi)^2 \geq 0$, і $\varepsilon > 0$. Тому за нерівністю Маркова (першою нерівністю Чебишова) одержуємо:

$$P\{\eta \geq \varepsilon^2\} \leq \frac{\mathbb{E}\eta}{\varepsilon^2}.$$

Отже,

$$P\{\eta \geq \varepsilon^2\} = P\{(\xi - \mathbb{E}\xi)^2 \geq \varepsilon^2\} = P\{|\xi - \mathbb{E}\xi| \geq \varepsilon\},$$

але $\mathbb{E}\eta = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = D\xi$. Звідси випливає, що

$$0 \leq P\{|\xi - \mathbb{E}\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}. \quad \blacktriangleright$$

Зрозуміло, що доведену нерівність можна записати і в такому вигляді:

$$P\{|\xi - \mathbb{E}\xi| < \varepsilon\} \geq 1 - \frac{D\xi}{\varepsilon^2}.$$

Нерівність Чебишова застосовується при побудові різних оцінок, при доведенні теорем. Крім цього, з її допомогою можна грубо оцінити ймовірність відхилення випадкової величини від свого математичного сподівання (для будь-якої випадкової величини).

п. 2.3. Характеристики положення і форми розподілу випадкової величини

Нехай ξ – деяка випадкова величина з функцією розподілу $F_\xi(x)$.

Означення 5. Розв'язок x_p нерівності

$$F_\xi(x) \leq p \leq F_\xi(x+0),$$

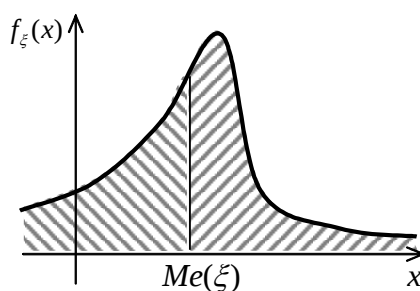
де число $p \in (0, 1)$ називається **квантиллю** порядку p .

Для неперервної випадкової величини записана нерівність фактично є рівнянням $F_\xi(x) = p$.

Означення 6. Медіаною розподілу випадкової величини ξ називається квантиль порядку $\frac{1}{2}$.

Крім позначення $x_{\frac{1}{2}}$, використовується ще й позначення $Me(\xi)$.

Враховуючи означення функції розподілу, одержуємо, що $P\{\xi < Me(\xi)\} = P\{\xi \geq Me(\xi)\} = \frac{1}{2}$. Таким чином, медіана характеризує „центральне значення” випадкової величини у тому розумінні, що ймовірність набути значення, менше за медіану, і ймовірність набути значення, більше за медіану, однакові.

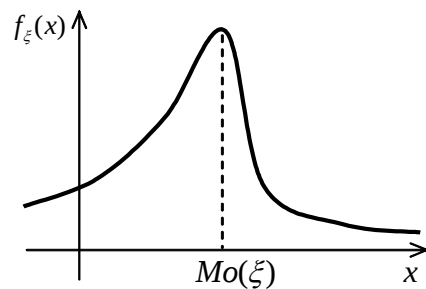
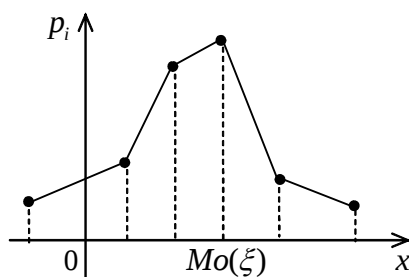


Означення 7. *Модою розподілу дискретної випадкової величини ξ називається найбільш імовірне її значення.*

Нехай неперервна випадкова величина ξ має щільність розподілу $f_\xi(x)$.

Означення 7'. *Модою розподілу неперервної випадкової величини ξ називається кожне дійсне значення x , на якому щільність розподілу $f_\xi(x)$ досягає максимуму.*

Моду розподілу випадкової величини ξ прийнято позначати $Mo(\xi)$. Розподіл, у якого мода єдина, називають **унімодальним** (наприклад, нормальний розподіл), у протилежному випадку – **полімодальним** (наприклад, рівномірний на відрізку розподіл). Глобальний максимум щільності розподілу випадкової величини називають модальний максимум або модус.



Використовуючи означення початкового і центрального моментів, а також математичного сподівання і дисперсії та їх властивості, можна просто одержати, що початковий момент першого порядку – це математичне сподівання:

$$\nu_1 = \mathbb{E} \xi,$$

центральный момент другого порядку – це дисперсія:

$$\mu_2 = D\xi,$$

а також можна виразити центральні моменти через початкові:

$$\mu_0 = 1,$$

$$\mu_1 = 0,$$

$$\mu_2 = \nu_2 - \nu_1^2,$$

$$\mu_3 = \mathbb{E}(\xi - \nu_1)^3 = \mathbb{E}(\xi^3 - 3\xi^2\nu_1 + 3\xi\nu_1^2 - \nu_1^3) =$$

$$= \nu_3 - 3\nu_2\nu_1 + 3\nu_1\nu_1^2 - \nu_1^3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3,$$

$$\mu_4 = \mathbb{E}(\xi - \nu_1)^4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 4\nu_1\nu_1^3 + \nu_1^4 =$$

$$= \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

Моменти вищих порядків застосовуються для більш детального опису розподілів.

Третій центральний момент служить для характеристики **асиметрії** (або „скошеності”) розподілу. Якщо розподіл симетричний відносно математичного сподівання (або, у механічній інтерпретації, маса розподілена симетрично відносно центра ваги), то всі центральні моменти непарного порядку (якщо вони існують) дорівнюють нулю. Дійсно, для дискретної випадкової величини в сумі

$$\mu_k = \mathbb{E}(\xi - \mathbb{E}\xi)^k = \sum_{i=1}^n (x_i - \mathbb{E}\xi)^k p_i$$

при симетричному відносно $\mathbb{E}\xi$ законі розподілу і непарному k кожному додатному доданку відповідає рівний йому за модулем від’ємний доданок, так що вся сума дорівнює нулю. Те ж саме, очевидно, справедливо і для неперервних випадкових величин: інтеграл

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

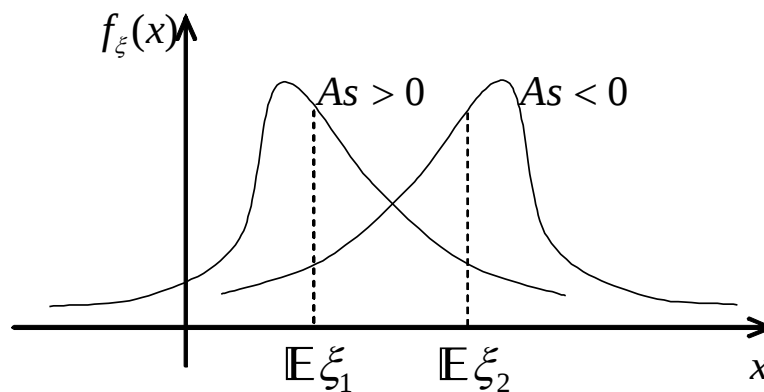
дорівнює нулю, як інтеграл від непарної функції у симетричних межах.

Тому природно як характеристику асиметрії розподілу вибрати якийсь з непарних моментів. Найпростіший з них – це третій центральний момент. Він має розмірність куба випадкової величини; щоб одержати безрозмірну характеристику, третій момент μ_3 ділять на куб середнього квадратичного відхилення σ .

Означення 8. Коефіцієнтом асиметрії, або просто *асиметрією*, випадкової величини ξ називається число As (якщо воно існує), яке дорівнює

$$As = \frac{\mu_3}{\sigma^3}.$$

На рисунку показано два асиметричних розподіли – один з додатною, а другий з від’ємною асиметріями.

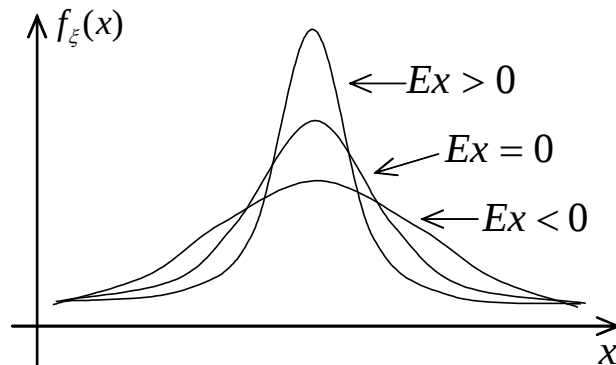


Четвертий центральний момент служить для характеристики так званої „крутості”, тобто гостровершинності розподілу.

Означення 9. Екссесом випадкової величини ξ називається число Ex (якщо воно існує), яке дорівнює

$$Ex = \frac{\mu_4}{\sigma^4} - 3.$$

Число 3 віднімається від частки $\frac{\mu_4}{\sigma^4}$ тому, що для нормального розподілу, як легко встановити, $\frac{\mu_4}{\sigma^4} = 3$. Отже, для нормально розподіленої випадкової величини ексцес дорівнює нулю. Якщо розподіл імовірностей випадкової величини ξ унімодальний і щільність розподілу



$f_\xi(x)$ має гострішу вершину, ніж щільність розподілу нормальної випадкової величини з такою самою дисперсією, то випадкова величина ξ володіє додатнім ексцесом ($Ex > 0$), якщо ж $f_\xi(x)$ менш гостроверха і більш згладжена у порівнянні зі щільністю відповідного нормального розподілу, то $Ex < 0$.

п. 2.4. Кореляція двох випадкових величин

Нехай задано дві випадкові величини ξ і η , для яких існують математичні сподівання і дисперсії $\mathbb{E}\xi$, $D\xi$, $\mathbb{E}\eta$, $D\eta$ причому $D\xi \neq 0$, $D\eta \neq 0$.

Означення 10. Випадкова величина $\overset{\circ}{\xi} = \xi - \mathbb{E}\xi$ називається *центрованою випадковою величиною* для ξ .

Вона характерна тим, що її математичне сподівання дорівнює нулю: $\mathbb{E}\overset{\circ}{\xi} = 0$. Аналогічно можна ввести центровану випадкову величину для η : $\overset{\circ}{\eta} = \eta - \mathbb{E}\eta$ і, відповідно, $\mathbb{E}\overset{\circ}{\eta} = 0$.

Якщо відомо $\overset{\circ}{\xi}$ і $\mathbb{E}\xi$, то можна перейти до ξ і навпаки.

Означення 11. Випадкова величина $\bar{\xi} = \frac{\overset{\circ}{\xi}}{\sqrt{D\xi}} = \frac{\overset{\circ}{\xi}}{\sigma_{\xi}}$ називається *стандартизованою (стандартною) випадковою величиною*.

Вона характерна тим, що її математичне сподівання також дорівнює нулю: $\mathbb{E}\bar{\xi} = 0$. Знайдемо дисперсію стандартизованої випадкової величини.

$$\begin{aligned} D\bar{\xi} &= \mathbb{E}(\bar{\xi} - \mathbb{E}\bar{\xi})^2 = \mathbb{E}\bar{\xi}^2 = \mathbb{E}\left(\frac{\xi - \mathbb{E}\xi}{\sigma_{\xi}}\right)^2 = \\ &= \frac{1}{\sigma_{\xi}^2} \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \frac{D\xi}{\sigma_{\xi}^2} = \frac{\sigma_{\xi}^2}{\sigma_{\xi}^2} = 1, \end{aligned}$$

тобто $D\bar{\xi} = 1$ – дисперсія стандартизованої випадкової величини дорівнює одиниці.

Аналогічно можна ввести $\bar{\eta} = \frac{\overset{\circ}{\eta}}{\sigma_{\eta}}$ і, відповідно, $\mathbb{E}\bar{\eta} = 0$, $D\bar{\eta} = 1$.

Завжди від будь-яких випадкових величин можна перейти до таких, у яких математичне сподівання дорівнює нулю, а дисперсія дорівнює одиниці.

Означення 12. Коваріацією випадкових величин ξ та η називається математичне сподівання добутку відповідних центрованих випадкових величин $\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}$:

$$\text{cov}(\xi, \eta) = \mathbb{E}\left(\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}\right) = \mathbb{E}\left((\xi - \mathbb{E}\xi) \cdot (\eta - \mathbb{E}\eta)\right).$$

Якщо ξ та η незалежні, то і $\overset{\circ}{\xi}$ та $\overset{\circ}{\eta}$ – незалежні, а це означає, що $\mathbb{E}\left(\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}\right) = \mathbb{E}\overset{\circ}{\xi} \cdot \mathbb{E}\overset{\circ}{\eta} = 0$, тобто $\text{cov}(\xi, \eta) = 0$ для незалежних випадкових величин. Отже, коваріація може розцінюватись як міра (характеристика) залежності двох випадкових величин.

Означення 13. Коваріація стандартизованих випадкових величин

$$\text{cov}\left(\bar{\xi}, \bar{\eta}\right) = \rho(\xi, \eta)$$

називається коефіцієнтом кореляції випадкових величин ξ та η .

Тобто коефіцієнт кореляції

$$\begin{aligned} \rho(\xi, \eta) &= \mathbb{E}\left(\left(\bar{\xi} - \mathbb{E}\bar{\xi}\right) \cdot \left(\bar{\eta} - \mathbb{E}\bar{\eta}\right)\right) = \\ &= \mathbb{E}\left(\bar{\xi} \cdot \bar{\eta}\right) = \mathbb{E}\left(\frac{\xi - \mathbb{E}\xi}{\sigma_{\xi}} \cdot \frac{\eta - \mathbb{E}\eta}{\sigma_{\eta}}\right) = \\ &= \frac{1}{\sigma_{\xi} \cdot \sigma_{\eta}} \mathbb{E}\left((\xi - \mathbb{E}\xi) \cdot (\eta - \mathbb{E}\eta)\right) = \frac{1}{\sigma_{\xi} \cdot \sigma_{\eta}} \cdot \text{cov}(\xi, \eta). \end{aligned}$$

Як же коефіцієнт кореляції характеризує залежність між випадковими величинами?

1. Якщо ξ та η незалежні, то $\rho(\xi, \eta) = 0$. Обернене твердження, взагалі кажучи, хибне (якщо $\rho(\xi, \eta) = 0$, то ξ та η не обов'язково незалежні).
2. Коефіцієнт кореляції для будь-яких ξ та η за модулем не перевищує одиниці: $|\rho(\xi, \eta)| \leq 1$.

◀ Розглянемо випадкову величину $\bar{\xi} \pm \bar{\eta}$. Для неї

$$\begin{aligned} 0 \leq D(\bar{\xi} \pm \bar{\eta}) &= \mathbb{E}((\bar{\xi} \pm \bar{\eta}) - \mathbb{E}(\bar{\xi} \pm \bar{\eta}))^2 = \mathbb{E}(\bar{\xi} \pm \bar{\eta})^2 = \\ &= \mathbb{E}(\bar{\xi}^2 + \bar{\eta}^2 \pm 2\bar{\xi}\bar{\eta}) = \mathbb{E}\bar{\xi}^2 + \mathbb{E}\bar{\eta}^2 \pm 2\mathbb{E}\bar{\xi}\bar{\eta} = \\ &= D\bar{\xi} + D\bar{\eta} \pm 2\rho(\xi, \eta) = 2(1 \pm \rho(\xi, \eta)). \end{aligned}$$

Звідси $1 + \rho(\xi, \eta) \geq 0$. Тому $\rho(\xi, \eta) \geq -1$. А також $1 - \rho(\xi, \eta) \geq 0$, тому $\rho(\xi, \eta) \leq 1$. Отже, $-1 \leq \rho(\xi, \eta) \leq 1$. ►

3. $|\rho(\xi, \eta)| = 1$ тоді і тільки тоді, коли існують такі дійсні числа A та B , що $\eta = A\xi + B$ (тобто коли зв'язок між випадковими величинами лінійний. Такий зв'язок виявиться пізніше найтіснішим зв'язком між випадковими величинами). Більш загально ця характеристика формулюється так: $|\rho(\xi, \eta)| = 1$ тоді і тільки тоді, коли існують такі дійсні числа A та B , що $P\{\eta = A\xi + B\} = 1$.

◀ 1) Нехай існують дійсні числа $A, B \in \mathbb{R}$, такі, що $P\{\eta = A\xi + B\} = 1$. Це означає, що ймовірність $P\{\eta - A\xi - B = 0\} = 1$. Звідси $\mathbb{E}(\eta - A\xi - B) = 0$ або $\mathbb{E}\eta = A \cdot \mathbb{E}\xi + B$. У простому першому формулюванні відразу $\mathbb{E}\eta = \mathbb{E}(A\xi + B) = A \cdot \mathbb{E}\xi + B$. Крім цього, знаємо, що $D\eta = A^2 D\xi$. Таким чином, якщо задано σ_ξ то $\sigma_\eta = |A| \sigma_\xi$. Отже,

$$\begin{aligned}\rho(\xi, \eta) &= \mathbb{E} \left(\frac{\xi - \mathbb{E}\xi}{\sigma_\xi} \cdot \frac{\eta - \mathbb{E}\eta}{\sigma_\eta} \right) = \mathbb{E} \left(\frac{\xi - \mathbb{E}\xi}{\sigma_\xi} \cdot \frac{A\xi + B - A \cdot \mathbb{E}\xi - B}{|A|\sigma_\xi} \right) = \\ &= \frac{A}{|A|\sigma_\xi^2} (\xi - \mathbb{E}\xi)^2 = \frac{A \cdot D\xi}{|A|\sigma_\xi^2} = \frac{A\sigma_\xi^2}{|A|\sigma_\xi^2} = \frac{A}{|A|}.\end{aligned}$$

Тобто $\rho(\xi, \eta) = 1$ при $A > 0$ і $\rho(\xi, \eta) = -1$ при $A < 0$. Те, що $A \neq 0$ означає, що жодна з випадкових величин не стала.

2) Нехай $|\rho(\xi, \eta)| = 1$. Припустимо, що $\rho(\xi, \eta) = 1$. Розглянемо випадкову величину $\bar{\xi} - \bar{\eta}$. Для неї знайдемо дисперсію:

$$\begin{aligned}0 \leq D(\bar{\xi} - \bar{\eta}) &= \mathbb{E}(\bar{\xi} - \bar{\eta})^2 = \mathbb{E}\bar{\xi}^2 + \mathbb{E}\bar{\eta}^2 - 2\mathbb{E}\bar{\xi}\bar{\eta} = \\ &= D\bar{\xi} + D\bar{\eta} - 2\rho(\xi, \eta) = 1 + 1 - 2 = 0.\end{aligned}$$

Отже, $D(\bar{\xi} - \bar{\eta}) = 0$. Тоді, за властивістю **1⁰** дисперсії, існує така стала C , що $P\{\bar{\xi} - \bar{\eta} = C\} = 1$, тобто з імовірністю одиниця

$$\frac{\xi - \mathbb{E}\xi}{\sigma_\xi} - \frac{\eta - \mathbb{E}\eta}{\sigma_\eta} = C.$$

А це й означає, що зв'язок між ξ та η — лінійний.

Аналогічно розглянемо випадок $\rho(\xi, \eta) = -1$, тоді одержимо

$$\frac{\xi - \mathbb{E}\xi}{\sigma_\xi} + \frac{\eta - \mathbb{E}\eta}{\sigma_\eta} = C \text{ і знову зв'язок лінійний. } \blacktriangleright$$

Якщо $\rho(\xi, \eta) > 0$, то **зв'язок** між ξ та η називається **прямим** (з ростом ξ зростає η і навпаки). Якщо $\rho(\xi, \eta) < 0$, то **зв'язок** між ξ та η називається **оберненим** (у тому розумінні, що при зростанні ξ випадкова величина η спадає і навпаки).

Приклад 1. Для випадкової величини $\eta = 3\xi - 4$ коефіцієнт кореляції $\rho(\xi, \eta) = 1$, бо лінійний зв'язок.

Приклад 2. Якщо $\eta = -27\xi + 17$, то $\rho(\xi, \eta) = -1$.

Розглянемо випадковий вектор $\eta = (\xi_1, \xi_2, \dots, \xi_n)$. Як і для скалярних випадкових величин, аналогічно можна ввести наступні поняття.

Означення 14. Математичним сподіванням випадкового вектора η називається

$$\mathbb{E} \eta = (\mathbb{E} \xi_1, \mathbb{E} \xi_2, \dots, \mathbb{E} \xi_n).$$

Означення 15. Змішаним початковим моментом порядку k випадкового вектора η називається математичне сподівання

$$\mathbb{E}(\xi_1^{\alpha_1} \xi_2^{\alpha_2} \dots \xi_n^{\alpha_n}), \text{ де } \alpha_i - \text{цілі невід'ємні числа, причому } \sum_{i=1}^n \alpha_i = k.$$

Означення 16. Змішаним центральним моментом порядку k випадкового вектора η називається математичне сподівання

$$\mathbb{E}((\xi_1 - \mathbb{E} \xi_1)^{\alpha_1} (\xi_2 - \mathbb{E} \xi_2)^{\alpha_2} \dots (\xi_n - \mathbb{E} \xi_n)^{\alpha_n}).$$

Найчастіше використовуються моменти порядку 1 і 2 (коли $k=1$ і $k=2$). При $k=1$ в означенні 15 одержуємо n початкових моментів, які утворюють вектор математичного сподівання. При $k=2$ утворюється n^2 моментів (відповідно, початкових або центральних).

Означення 17. Матриця $\text{Cov}(\xi_1, \xi_2, \dots, \xi_n)$, складена зі змішаних центральних моментів другого порядку, називається матрицею коваріації випадкового вектора η .

Ці центральні моменти другого порядку мають такий вигляд:

$$\mathbb{E}(\xi_i - \mathbb{E} \xi_i)^{\alpha_i} (\xi_j - \mathbb{E} \xi_j)^{\alpha_j} = \mu_{ij}, \quad i, j = \overline{1, n}, \quad \alpha_i + \alpha_j = 2, \quad \alpha_i, \alpha_j \in \{0, 1, 2\}.$$

$$\text{Тоді, за означенням, } \text{Cov}(\xi_1, \xi_2, \dots, \xi_n) = \|\mu_{ij}\|_{i,j=1}^n = \|\text{cov}(\xi_i, \xi_j)\|_{i,j=1}^n.$$

На діагоналі цієї матриці розміщені числа μ_{ii} , які дорівнюють: $\mu_{ii} = D\xi_i$, а інші елементи матриці: $\mu_{ij} = \text{cov}(\xi_i, \xi_j)$, $i \neq j$.

Означення 18. Кореляційною матрицею $\mathbb{R}(\xi_1, \xi_2, \dots, \xi_n)$ називається матриця, елементами якої є коефіцієнти кореляції

$$\mathbb{R}(\xi_1, \xi_2, \dots, \xi_n) = \left\| \rho(\xi_i, \xi_j) \right\|_{i,j=1}^n.$$

Існує простий зв'язок між коваріаційною і кореляційною матрицями. Запишемо його для матриці 2×2 :

$$\begin{aligned} \mathbb{R}(\xi_1, \xi_2) &= \begin{pmatrix} 1 & \frac{\text{cov}(\xi_1, \xi_2)}{\sigma_{\xi_1} \sigma_{\xi_2}} \\ \frac{\text{cov}(\xi_1, \xi_2)}{\sigma_{\xi_1} \sigma_{\xi_2}} & 1 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{1}{\sigma_{\xi_1}} & 0 \\ 0 & \frac{1}{\sigma_{\xi_2}} \end{pmatrix} \begin{pmatrix} \text{cov}(\xi_1, \xi_1) & \text{cov}(\xi_1, \xi_2) \\ \text{cov}(\xi_2, \xi_1) & \text{cov}(\xi_2, \xi_2) \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_{\xi_1}} & 0 \\ 0 & \frac{1}{\sigma_{\xi_2}} \end{pmatrix}. \end{aligned}$$

Кажуть, що **випадковий вектор** (ξ_1, ξ_2) **розподілений за законом Гаусса**, якщо щільність сумісного розподілу $f_{\xi_1 \xi_2}(x_1, x_2)$ має такий вигляд:

$$\begin{aligned} f_{\xi_1 \xi_2}(x_1, x_2) &= \frac{1}{2\pi \sigma_{\xi_1} \sigma_{\xi_2} \sqrt{1-r^2}} \times \\ &\times \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{(x_1-a)^2}{\sigma_{\xi_1}^2} - 2r \frac{(x_1-a)(x_2-b)}{\sigma_{\xi_1} \sigma_{\xi_2}} + \frac{(x_2-b)^2}{\sigma_{\xi_2}^2} \right] \right\}, \end{aligned}$$

числа $a, b, \sigma_{\xi_1}, \sigma_{\xi_2}, r$ називаються **параметрами розподілу**. Легко підрахувати, що

$$\mathbb{E} \xi_1 = a, \quad \mathbb{E} \xi_2 = b, \quad D\xi_1 = \sigma_{\xi_1}^2, \quad D\xi_2 = \sigma_{\xi_2}^2, \quad \rho(\xi_1, \xi_2) = r.$$

Якщо випадкові величини ξ_1 та ξ_2 незалежні, то $r = 0$. У цьому випадку щільність сумісного розподілу $f_{\xi_1\xi_2}(x_1, x_2)$ (яка дорівнює добутку щільностей) записується так:

$$f_{\xi_1\xi_2}(x_1, x_2) = \frac{1}{\sigma_{\xi_1}\sqrt{2\pi}} e^{-\frac{(x_1-a)^2}{2\sigma_{\xi_1}^2}} \frac{1}{\sigma_{\xi_2}\sqrt{2\pi}} e^{-\frac{(x_2-b)^2}{2\sigma_{\xi_2}^2}}.$$

Висновок: для нормально розподілених випадкових величин рівність нулю коефіцієнта кореляції є достатньою умовою незалежності випадкових величин. Це видно з того, що щільність сумісного розподілу дорівнює добутку одновимірних щільностей. У цьому випадку матриця коваріації (для нормально розподілених випадкових величин ξ_1, ξ_2) має такий вигляд:

$$\text{Cov}(\xi_1, \xi_2) = \begin{pmatrix} D\xi_1 & r\sigma_{\xi_1}\sigma_{\xi_2} \\ r\sigma_{\xi_1}\sigma_{\xi_2} & D\xi_2 \end{pmatrix},$$

а кореляційна матриця

$$\mathbb{R}(\xi_1, \xi_2) = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}.$$

Якщо випадкову величину вивчають, використовуючи тільки перший початковий момент (математичне сподівання) і другий центральний момент (дисперсію), то кажуть, що випадкова величина вивчається у широкому розумінні. Якщо ж вивчення ведеться на основі відомого закону розподілу, то кажуть, що випадкова величина вивчається у вузькому розумінні.

§3. Характеристична функція випадкової величини

Нехай на ймовірнісному просторі $(\Omega, \mathcal{F}, P)$ задана деяка випадкова величина ξ , тобто задана її функція розподілу або щільність.

Означення 1. *Характеристичною функцією випадкової величини ξ називається функція*

$$\varphi_{\xi}(t) \stackrel{\text{def}}{=} \mathbb{E} e^{it\xi},$$

де $t \in \mathbb{R}$ – параметр.

Це комплексно-значна функція дійсного аргументу. Можна записати $\varphi_{\xi}(t)$ через інтеграл Стільтьєса у такому вигляді:

$$\varphi_{\xi}(t) = \int_{-\infty}^{\infty} e^{itx} dF_{\xi}(x),$$

маючи на увазі, що такий інтеграл збігається. Якщо вважати, що випадкова величина абсолютно неперервного типу, то

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

З точністю до множника – це перетворення Фур'є^{*} щільності розподілу.

Використовуючи формулу Ейлера^{**}, $e^{ix} = \cos(x) + i \sin(x)$, можна розкласти характеристичну функцію на дійсну та уявну частини. Оскільки

$$e^{it\xi} = \cos(t\xi) + i \sin(t\xi),$$

^{*} Жан Батіст Жозеф Фур'є (Jean Baptiste Joseph Fourier, 1768-1830) – французький математик і фізик.

^{**} Леонард Ейлер (Leonhard Euler, 1707-1783) – швейцарський математик та фізик.

то характеристична функція $\varphi_\xi(t)$ записується так:

$$\varphi_\xi(t) = \mathbb{E}[\cos(t\xi) + i\sin(t\xi)] = \mathbb{E}[\cos(t\xi)] + i \cdot \mathbb{E}[\sin(t\xi)],$$

тобто її дійсна частина дорівнює $\mathbb{E}[\cos(t\xi)]$, а уявна частина дорівнює $\mathbb{E}[\sin(t\xi)]$.

Властивості характеристичної функції

1⁰. Для будь-якої випадкової величини ξ характеристична функція $\varphi_\xi(t)$ існує при кожному $t \in \mathbb{R}$.

◀ Досить показати, що інтеграл в означенні завжди збігається. Дійсно:

$$\left| \int_{-\infty}^{\infty} e^{itx} dF_\xi(x) \right| \leq \int_{-\infty}^{\infty} |e^{itx}| dF_\xi(x) = \int_{-\infty}^{\infty} dF_\xi(x) = 1,$$

оскільки модуль функції e^{itx} дорівнює 1, тобто $|e^{itx}| = 1$. ►

2⁰. Характеристична функція $\varphi_\xi(t)$ випадкової величини ξ

а) рівномірно неперервна на числовій осі,

б) $|\varphi_\xi(t)| \leq 1$,

в) $\varphi_\xi(0) = 1$.

◀ б) і в) елементарно перевіряються і доведені в **1⁰**. Доведемо а): що функція $\varphi_\xi(t)$ рівномірно неперервна на числовій осі. За означенням рівномірної неперервності, потрібно показати, що для будь-якого $\varepsilon > 0$ існує h незалежне від t , таке, що $|\varphi_\xi(t+h) - \varphi_\xi(t)| < \varepsilon$. Для $t \in \mathbb{R}$, $h > 0$ розглянемо різницю

$$\varphi_{\xi}(t+h) - \varphi_{\xi}(t) = \int_{-\infty}^{\infty} e^{itx} (e^{ihx} - 1) dF_{\xi}(x),$$

і оцінимо її модуль. Маємо:

$$\begin{aligned} |\varphi_{\xi}(t+h) - \varphi_{\xi}(t)| &\leq \int_{-\infty}^{\infty} |e^{itx} (e^{ihx} - 1)| dF_{\xi}(x) \leq \int_{-\infty}^{\infty} |e^{ihx} - 1| dF_{\xi}(x) = \\ &= \int_{|x| \leq A} |e^{ihx} - 1| dF_{\xi}(x) + \int_{|x| > A} |e^{ihx} - 1| dF_{\xi}(x) \end{aligned}$$

для довільного $A > 0$. Оскільки $|a - b| \leq |a| + |b|$ і $|e^{ihx}| = 1$, то $|e^{ihx} - 1| \leq 2$, тому

$$|\varphi_{\xi}(t+h) - \varphi_{\xi}(t)| \leq \int_{|x| \leq A} |e^{ihx} - 1| dF_{\xi}(x) + 2 \int_{|x| > A} dF_{\xi}(x).$$

Візьмемо довільне $\varepsilon > 0$; виберемо A настільки великим, щоб $\int_{|x| > A} dF_{\xi}(x) < \frac{\varepsilon}{4}$, і підберемо настільки мале h , щоб для $|x| \leq A$

виконувалась нерівність $|e^{ihx} - 1| < \frac{\varepsilon}{2}$. Тоді

$$|\varphi_{\xi}(t+h) - \varphi_{\xi}(t)| < \frac{\varepsilon}{2} \int_{|x| \leq A} dF_{\xi}(x) + \frac{\varepsilon}{2} \leq \varepsilon.$$

Тобто для будь-якого $\varepsilon > 0$ можна знайти таке h , незалежне від t , що

$$|\varphi_{\xi}(t+h) - \varphi_{\xi}(t)| < \varepsilon,$$

а це і треба було довести. ►

З⁰. Характеристична функція суми двох незалежних випадкових величин дорівнює добутку їх характеристичних функцій: якщо

випадкові величини ξ_1 і ξ_2 незалежні, а $\varphi_{\xi_1}(t)$, $\varphi_{\xi_2}(t)$ – їх характеристичні функції, то характеристична функція

$$\varphi_{\xi_1+\xi_2}(t) = \varphi_{\xi_1}(t)\varphi_{\xi_2}(t).$$

◀ Якщо ξ_1 і ξ_2 незалежні, то $e^{it\xi_1}$, $e^{it\xi_2}$ також незалежні. Знайдемо $\varphi_{\xi_1+\xi_2}$:

$$\varphi_{\xi_1+\xi_2} = \mathbb{E}e^{it(\xi_1+\xi_2)} = \mathbb{E}(e^{it\xi_1}e^{it\xi_2}) = \mathbb{E}e^{it\xi_1} \cdot \mathbb{E}e^{it\xi_2} = \varphi_{\xi_1}(t) \cdot \varphi_{\xi_2}(t). \quad \blacktriangleright$$

Цю властивість можна узагальнити: якщо випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ незалежні у сукупності, то характеристична функція суми дорівнює добутку характеристичних функцій. Цієї вимоги (незалежності у сукупності) навіть забагато. Досить вважати, що незалежними є ξ_i та $\xi_1 + \xi_2 + \xi_3 + \dots + \xi_{i-1}$, $i = \overline{1, n}$ (щоб одна з випадкових величин була незалежна від суми попередніх), тоді

$$\varphi_{\sum_{i=1}^n \xi_i}(t) = \prod_{i=1}^n \varphi_{\xi_i}(t).$$

Характеристичні функції зручно використовувати як апарат для дослідження випадкових величин.

4⁰. Нехай випадкова величина $\eta = a\xi + b$, де a і b – постійні. Тоді характеристична функція $\varphi_{\eta}(t)$ випадкової величини η дорівнює

$$\varphi_{\eta}(t) = \varphi_{a\xi+b}(t) = e^{itb} \varphi_{\xi}(at).$$

$$\blacktriangleleft \varphi_{\eta}(t) = \mathbb{E}e^{it\eta} = \mathbb{E}e^{it(a\xi+b)} = \mathbb{E}e^{itb} e^{i(at)\xi} = e^{itb} \mathbb{E}e^{i(at)\xi} = e^{itb} \varphi_{\xi}(at). \quad \blacktriangleright$$

5⁰. Якщо існує абсолютний момент порядку k випадкової величини ξ $\mathbb{E}|\xi|^k$, то характеристична функція $\varphi_{\xi}(t)$ диференційовна k разів і похідна k -го порядку в нулі дорівнює $\varphi_{\xi}^{(k)}(0) = i^k \cdot \mathbb{E}\xi^k$.

(Тут $\mathbb{E}\xi^k$ – початковий момент порядку k випадкової величини ξ).

◀ За означенням, $\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \int_{-\infty}^{\infty} e^{itx} dF_\xi(x)$. Формально продиференціюємо під знаком інтеграла

$$\varphi'_\xi(t) = i \int_{-\infty}^{\infty} x e^{itx} dF_\xi(x)$$

– цей інтеграл існує, бо

$$\left| \int_{-\infty}^{\infty} x e^{itx} dF_\xi(x) \right| \leq \int_{-\infty}^{\infty} |x| dF_\xi(x),$$

а тому, за умовою, обмежений. Звідси випливає законність диференціювання, тобто $\varphi_\xi(t)$ – диференційовна. Тоді

$$\varphi_\xi^{(k)}(t) = i^k \int_{-\infty}^{\infty} x^k e^{itx} dF_\xi(x), \text{ тому } \varphi_\xi^{(k)}(0) = i^k \int_{-\infty}^{\infty} x^k dF_\xi(x) = i^k \cdot \mathbb{E}\xi^k. \quad \blacktriangleright$$

Використаємо результат властивості 5⁰: $\varphi'_\xi(0) = i \cdot \mathbb{E}\xi$. Звідси $\mathbb{E}\xi = -i\varphi'_\xi(0)$. У лівій частині цієї рівності математичне сподівання є дійсним числом, тому у правій частині рівності також має бути дійсне число. Отже, $\varphi'_\xi(0)$ – чисто уявне. Далі: $\varphi''_\xi(0) = i^2 \cdot \mathbb{E}\xi^2 = -\mathbb{E}\xi^2$, але $D\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2$. Тоді $D\xi = -\varphi''_\xi(0) + [\varphi'_\xi(0)]^2$, причому $D\xi \geq 0$, а тому $[\varphi'_\xi(0)]^2 \geq \varphi''_\xi(0)$.

Сформулюємо ще ряд властивостей без доведення.

6⁰. За відомою характеристичною функцією можна знайти функцію розподілу випадкової величини. Для цього потрібно використати обернене перетворення Фур'є і функцію розподілу знайти у вигляді границі.

7⁰. Якщо послідовність функцій розподілу $F_{\xi_1}(x), F_{\xi_2}(x), \dots, F_{\xi_n}(x), \dots$ збігається, в основному, до деякої функції розподілу $F_{\xi}(x)$ (тобто збігається у точках неперервності функції $F_{\xi}(x)$), то послідовність $\varphi_{\xi_1}(t), \varphi_{\xi_2}(t), \dots, \varphi_{\xi_n}(t), \dots$ збігається до характеристичної функції $\varphi_{\xi}(t)$.

8⁰. Якщо послідовність характеристичних функцій $\varphi_{\xi_n}(t) \xrightarrow{n \rightarrow \infty} \varphi_{\xi}(t)$ (поточкова збіжність), то послідовність функцій розподілу $F_{\xi_n}(x) \xrightarrow{n \rightarrow \infty} F_{\xi}(x)$, причому збіжність рівномірна. (Ця властивість протилежна до 7⁰).

Звідси впливає ще одне застосування характеристичних функцій. Їх зручно використовувати при граничних переходах: якщо характеристичні функції збігаються, то можна робити висновок про функцію розподілу, до якої збігаються відповідні послідовності функцій розподілу.

Приклад 1. Знайти характеристичну функцію випадкової величини, заданої таким законом розподілу:

ξ	-1	1
P	0,5	0,5

$$\blacktriangleleft \varphi_{\xi}(t) = \mathbb{E}e^{-it\xi} = \frac{1}{2}e^{-it} + e^{it}\frac{1}{2} = \frac{e^{it} + e^{-it}}{2} = \cos t. \quad \blacktriangleright$$

Приклад 2. Знайти характеристичну функцію випадкової величини ξ , розподіленої за законом Пуассона з параметром λ .

$$\blacktriangleleft P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots \quad \text{Тоді}$$

$$\varphi_{\xi}(t) = \mathbb{E}e^{it\xi} = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}}.$$

Тобто для розподілу Пуассона характеристична функція така:

$$\varphi_{\xi}(t) = e^{\lambda(e^{it}-1)}. \quad \blacktriangleright$$

Приклад 3. Знайти характеристичну функцію випадкової величини ξ , розподіленої нормально з параметрами 0 і 1: $N(0, 1)$.

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

$$\varphi_{\xi}(t) = \mathbb{E}e^{it\xi} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx,$$

продиференціюємо цю рівність по t :

$$\begin{aligned} \varphi'_{\xi}(t) &= \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{itx} e^{-\frac{x^2}{2}} dx = -\frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} d\left(e^{-\frac{x^2}{2}}\right) = -\frac{i}{\sqrt{2\pi}} e^{itx} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \\ &+ \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} ite^{itx} dx = -t \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx, \end{aligned}$$

$$\text{але } \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx = \varphi_{\xi}(t). \text{ Тобто}$$

$$\varphi'_{\xi}(t) = -t\varphi_{\xi}(t),$$

$$\varphi_{\xi}(0) = 1.$$

Розв'язок такої задачі Коші $\varphi_{\xi}(t) = e^{-\frac{t^2}{2}}$ і є характеристичною функцією для випадкової величини $\xi \in N(0, 1)$. $\blacktriangleright$

Приклад 4. Припустимо, що випадкова величина ξ розподілена нормально з параметрами $N(a, \sigma^2)$, тобто $\mathbb{E}\xi = a$, $D\xi = \sigma^2$. Знайти її характеристичну функцію.

◀ Розглянемо стандартизовану для ξ випадкову величину $\eta = \frac{\xi - a}{\sigma}$. Ця випадкова величина також нормально розподілена, але з параметрами 0 і 1 $\eta \in N(0, 1)$ і тому $\varphi_\eta(t) = e^{-\frac{t^2}{2}}$, причому $\xi = \sigma\eta + a$ і за властивістю 4⁰:

$$\varphi_\xi(t) = e^{ita} \varphi_\eta(\sigma t) = e^{ita} e^{-\frac{(\sigma t)^2}{2}} = e^{ita - \frac{\sigma^2 t^2}{2}}.$$

Тобто характеристична функція нормально розподіленої випадкової величини ξ з параметрами $N(a, \sigma^2)$ записується так:

$$\varphi_\xi(t) = e^{ita - \frac{\sigma^2 t^2}{2}}. \quad \blacktriangleright$$

Якщо, наприклад, задано $\varphi_\xi(t) = e^{it - \frac{2t^2}{3}}$, то $E\xi = 1$, $\sigma = \frac{2}{\sqrt{3}}$, а тому легко виписати щільність розподілу

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{3}}{2} e^{-\frac{(x-1)^2 3}{2 \cdot 4}}.$$

Приклад 5. Нехай випадкова величина ξ_1 розподілена за законом Пуассона з параметром λ_1 , а випадкова величина ξ_2 – за законом Пуассона з параметром λ_2 . Вважатимемо, що ξ_1 і ξ_2 – незалежні. За яким законом розподілена випадкова величина $\eta = \xi_1 + \xi_2$?

◀ Згідно з прикладом 2: $\varphi_{\xi_1}(t) = e^{\lambda_1(e^{it}-1)}$ аналогічно $\varphi_{\xi_2}(t) = e^{\lambda_2(e^{it}-1)}$. За властивістю 3⁰:

$$\varphi_\eta(t) = \varphi_{\xi_1} \cdot \varphi_{\xi_2} = e^{(\lambda_1 + \lambda_2)(e^{it}-1)}.$$

За цією характеристичною функцією можна відновити функцію розподілу η . Видно, що це характеристична функція випадкової

величини, розподіленої за законом Пуассона з параметром $\lambda_1 + \lambda_2$, тобто

$$P\{\eta = k\} = \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)}, \quad k = 0, 1, 2, \dots \quad \blacktriangleright$$

Можна довести і такий факт: якщо випадкова величина η , що є сумою деяких випадкових величин, розподілена за законом Пуассона, то і доданки розподілені за законом Пуассона.

Приклад 6. Нехай ξ_1 – нормально розподілена випадкова величина з параметрами $N(a_1, \sigma_1^2)$, ξ_2 – нормально розподілена випадкова величина з параметрами $N(a_2, \sigma_2^2)$, причому ξ_1, ξ_2 – незалежні. Як розподілена $\eta = \xi_1 + \xi_2$?

◀ Згідно з прикладом 4, запишемо

$$\varphi_{\xi_1}(t) = e^{ita_1 - \frac{\sigma_1^2 t^2}{2}}, \quad \varphi_{\xi_2}(t) = e^{ita_2 - \frac{\sigma_2^2 t^2}{2}}.$$

Тоді, за властивістю 3⁰,

$$\varphi_{\eta}(t) = e^{it(a_1 + a_2) - \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}}.$$

Характеристична функція випадкової величини η має таку саму структуру, як і ξ_1, ξ_2 , а це означає, що η – розподілена нормально з параметрами

$$\mathbb{E}(\xi_1 + \xi_2) = a_1 + a_2, \quad D(\xi_1 + \xi_2) = \sigma_1^2 + \sigma_2^2, \quad \sigma(\xi_1 + \xi_2) = \sqrt{\sigma_1^2 + \sigma_2^2},$$

тобто $\eta \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$. $\blacktriangleright$

§4. Послідовності випадкових величин

Нехай на ймовірнісному просторі $(\Omega, \mathcal{F}, P)$ задано випадкові величини $\xi_1, \xi_2, \dots, \xi_n, \dots$, тобто послідовність випадкових величин $\{\xi_n\}_1^\infty$.

Означення 1. Кажуть, що послідовність випадкових величин $\{\xi_n\}_1^\infty$ збігається за ймовірністю до випадкової величини ξ , якщо для будь-якого $\varepsilon > 0$ границя

$$\lim_{n \rightarrow \infty} P\{\omega : |\xi_n - \xi| > \varepsilon\} = 0.$$

Останню рівність можна записати так:

$$\lim_{n \rightarrow \infty} P\{\omega : |\xi_n - \xi| \leq \varepsilon\} = 1.$$

Використовуються такі позначення для збіжності за ймовірністю:

$$p\lim_{n \rightarrow \infty} \xi_n = \xi \quad \text{або} \quad \xi_n \xrightarrow[p]{n \rightarrow \infty} \xi.$$

Властивості збіжності за ймовірністю

1⁰. Якщо $\xi_n \xrightarrow[p]{n \rightarrow \infty} \xi$, а $f(x)$ – функція, визначена і неперервна на $\mathbb{R}$, то

$$f(\xi_n) \xrightarrow[p]{n \rightarrow \infty} f(\xi).$$

◀ Неперервність функції $f(x)$ означає, що вона рівномірно неперервна на кожній замкненій множині, тому для будь-якого $\varepsilon > 0$ існує $\delta(\varepsilon) > 0$: що з умови $|\xi_n - \xi| \leq \delta$ випливає $|f(\xi_n) - f(\xi)| \leq \varepsilon$.

Виконання умови $|\xi_n - \xi| \leq \delta$ означає, що виконується подія

$\{\omega: |\xi_n - \xi| \leq \delta\}$, а з неї випливає виконання нерівності $|f(\xi_n) - f(\xi)| \leq \varepsilon$, тобто виконання події $\{\omega: |f(\xi_n) - f(\xi)| \leq \varepsilon\}$. Отже, перша подія викликає другу, тобто є включення подій:

$$\{\omega: |\xi_n - \xi| \leq \delta\} \subset \{\omega: |f(\xi_n) - f(\xi)| \leq \varepsilon\}.$$

Тому, за властивістю **4**⁰ ймовірності,

$$P\{|\xi_n - \xi| \leq \delta\} \leq P\{|f(\xi_n) - f(\xi)| \leq \varepsilon\} \leq 1,$$

але $P\{|\xi_n - \xi| \leq \delta\} = 1$. Тому $P\{|f(\xi_n) - f(\xi)| \leq \varepsilon\} = 1$. ►

2⁰. Нехай $\xi_n \xrightarrow{p} \xi$ та існує стала $C > 0$, така, що ймовірність $P\{|\xi_n| \leq C\} = 1$ (це означає обмеженість за ймовірністю 1, тобто майже всюди обмеженість). Тоді існує математичне сподівання $\mathbb{E}\xi$ і послідовність математичних сподівань ξ_n прямує до математичного сподівання випадкової величини ξ :

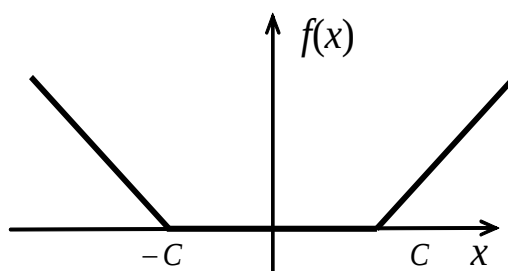
$$\mathbb{E}\xi_n \xrightarrow{n \rightarrow \infty} \mathbb{E}\xi.$$

◀ Із того, що $P\{|\xi_n| \leq C\} = 1$, випливає

$$|\mathbb{E}\xi_n| \leq \mathbb{E}|\xi_n| \leq \mathbb{E}C = C,$$

тобто $\mathbb{E}\xi_n$ існує. Доведемо, що існує $\mathbb{E}\xi$. Покажемо спочатку, що $P\{|\xi| \leq C\} = 1$. Нехай $f(x)$ – деяка неперервна функція, така, що

$$f(x) = \begin{cases} 0, & |x| \leq C, \\ > 0, & |x| > C, \end{cases} \quad \begin{cases} 0, & |x| \leq C, \\ -C, & |x| > C. \end{cases}$$

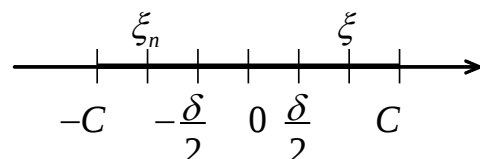


Оскільки за умовою $P\{|\xi_n| \leq C\} = 1$, то в цьому випадку ймовірність $P\{f(\xi_n) = 0\} = 1$. Але за властивістю $\mathbf{1}^0 \quad f(\xi_n) \xrightarrow[p]{n \rightarrow \infty} f(\xi)$. Тому з імовірністю 1 $f(\xi) = 0$, тобто $P\{f(\xi) = 0\} = 1$. Тоді, за означенням функції $f(x)$, одержуємо, що і $P\{|\xi| \leq C\} = 1$. Звідси, за властивостями математичного сподівання, маємо: $|\mathbb{E}\xi| \leq \mathbb{E}|\xi| \leq \mathbb{E}C = C$. Отже, математичне сподівання ξ існує.

Залишилося показати, що $\mathbb{E}\xi_n \xrightarrow[n \rightarrow \infty]{} \mathbb{E}\xi$. Розглянемо

$$|\mathbb{E}\xi_n - \mathbb{E}\xi| = |\mathbb{E}(\xi_n - \xi)| \leq \mathbb{E}|\xi_n - \xi|.$$

Оцінимо $\xi_n - \xi$. Для будь-якого достатньо малого $\delta > 0$ введемо у розгляд випадкову величину $\chi(\delta)$, яку називають індикаторною функцією:

$$\chi(\delta) = \begin{cases} 0, & \omega \in \{\omega: |\xi_n - \xi| \leq \delta\}, \\ 1, & \omega \in \{\omega: |\xi_n - \xi| > \delta\}. \end{cases}$$


Видно, що можна записати

$$|\xi_n - \xi| \leq \delta + 2C\chi(\delta).$$

Остання нерівність – це нерівність між двома випадковими величинами, тому така ж нерівність виконується і між їх математичними сподіваннями:

$$\begin{aligned} \mathbb{E}|\xi_n - \xi| &\leq \mathbb{E}(\delta + 2C\chi(\delta)) = \delta + 2C \cdot \mathbb{E}(\chi(\delta)) = \\ &\left| \begin{array}{l} \text{але } \chi(\delta) \text{ – дискретна випадкова величина з двома значеннями } 0 \text{ і } \\ 1, \text{ тоді} \end{array} \right| \\ &= \delta + 2C(0 \cdot P\{|\xi_n - \xi| \leq \delta\} + 1 \cdot P\{|\xi_n - \xi| > \delta\}) = \\ &= \delta + 2C \cdot P\{|\xi_n - \xi| > \delta\}. \end{aligned}$$

В одержаній нерівності

$$\mathbb{E}|\xi_n - \xi| \leq \delta + 2C \cdot P\{|\xi_n - \xi| > \delta\},$$

перейдемо до границі при $n \rightarrow \infty$. За умовою, $\lim_{n \rightarrow \infty} P\{|\xi_n - \xi| > \delta\} = 0$.

Тоді одержуємо: $\lim_{n \rightarrow \infty} \mathbb{E}(|\xi_n - \xi|) \leq \delta$. Тому і $\lim_{n \rightarrow \infty} |\mathbb{E}\xi_n - \mathbb{E}\xi| \leq \delta$, а це означає, що $\lim_{n \rightarrow \infty} \mathbb{E}\xi_n = \mathbb{E}\xi$. ►

Означення 2. Кажуть, що послідовність випадкових величин $\{\xi_n\}_1^\infty$ збігається в середньому до випадкової величини ξ , якщо існують математичні сподівання $\mathbb{E}\xi$, $\mathbb{E}\xi_n$ для всіх n , і границя

$$\lim_{n \rightarrow \infty} \mathbb{E}|\xi_n - \xi| = 0.$$

З⁰. Зі збіжності в середньому випливає збіжність за ймовірністю.

◀ Дійсно, для невід'ємної випадкової величини $|\xi_n - \xi|$ і будь-якого $\varepsilon > 0$ виконується нерівність Маркова (перша нерівність Чебишова):

$$P\{|\xi_n - \xi| > \varepsilon\} \leq \frac{1}{\varepsilon} \mathbb{E}|\xi_n - \xi|.$$

Якщо тепер перейти до границі при $n \rightarrow \infty$ і використати збіжність в середньому (тобто права частина прямує до 0), то з неї випливає, що і ліва частина прямує до 0 при $n \rightarrow \infty$, а це й означає, що $\{\xi_n\}_1^\infty$ збігається за ймовірністю до випадкової величини ξ . ►

Але зі збіжності за ймовірністю не завжди випливає збіжність в середньому.

Означення 3. Кажуть, що послідовність випадкових величин $\{\xi_n\}_1^\infty$ збігається при $n \rightarrow \infty$ до випадкової величини ξ в середньому квадратичному, якщо існують математичні сподівання $\mathbb{E}\xi_n^2$, $\mathbb{E}\xi^2$ і границя

$$\lim_{n \rightarrow \infty} \mathbb{E}|\xi_n - \xi|^2 = 0.$$

4⁰. Зі збіжності в середньому квадратичному випливає збіжність у середньому, а тому і збіжність за ймовірністю.

◀ Дійсно,

$$0 \leq D(|\xi_n - \xi|) = \mathbb{E}(|\xi_n - \xi|)^2 - (\mathbb{E}|\xi_n - \xi|)^2.$$

Звідси очевидно випливає, що

$$(\mathbb{E}|\xi_n - \xi|)^2 \leq \mathbb{E}(|\xi_n - \xi|)^2.$$

Добуваючи квадратний корінь, одержуємо

$$\mathbb{E}|\xi_n - \xi| \leq \sqrt{\mathbb{E}|\xi_n - \xi|^2},$$

і якщо є збіжність в середньому квадратичному (тобто права частина прямує до 0), то є збіжність і в середньому (ліва частина прямує до 0 при $n \rightarrow \infty$). ▶

Означення 4. Кажуть, що послідовність випадкових величин $\{\xi_n\}_1^\infty$ збігається при $n \rightarrow \infty$ до випадкової величини ξ **майже напевно** або **з ймовірністю одиниця**, якщо

$$P\left\{\omega: \lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)\right\} = 1.$$

Іншими словами, послідовність $\{\xi_n\}_1^\infty$ збігається до ξ майже напевно, якщо множина всіх $\omega \in \Omega$, для яких $\lim_{n \rightarrow \infty} \xi_n(\omega) \neq \xi(\omega)$, має ймовірність нуль.

Використовуються такі позначення для збіжності з ймовірністю одиниця:

$$\xi_n \xrightarrow[n \rightarrow \infty]{a.s.} \xi \quad \text{або} \quad \xi_n \rightarrow \xi \text{ майже напевно (almost sure).}$$

Найслабша збіжність за ймовірністю, потім збіжність у середньому, потім збіжність у середньому квадратичному і відтак збіжність з ймовірністю одиниця.

5⁰. Якщо $\xi_n \rightarrow \xi$ в середньому квадратичному, то

$$\mathbb{E} \xi_n \xrightarrow{n \rightarrow \infty} \mathbb{E} \xi \quad \text{і} \quad \mathbb{E} \xi_n^2 \xrightarrow{n \rightarrow \infty} \mathbb{E} \xi^2.$$

6⁰. Для того, щоб послідовність $\{\xi_n\}_1^\infty$ збігалася в середньому квадратичному до деякої сталої C , необхідно і досить, щоб

$$\lim_{n \rightarrow \infty} \mathbb{E} \xi_n = C, \quad \text{а} \quad \lim_{n \rightarrow \infty} D \xi_n = 0.$$

◀ Якщо $\xi_n \rightarrow C$ в середньому квадратичному, то, за властивістю 5⁰,

$$\mathbb{E} \xi_n \xrightarrow{n \rightarrow \infty} \mathbb{E} C = C, \quad \text{і} \quad \mathbb{E} \xi_n^2 \xrightarrow{n \rightarrow \infty} C^2.$$

Тому

$$D \xi_n = \mathbb{E} \xi_n^2 - (\mathbb{E} \xi_n)^2 \xrightarrow{n \rightarrow \infty} 0.$$

Навпаки, якщо $\lim_{n \rightarrow \infty} \mathbb{E} \xi_n = C$ і $\lim_{n \rightarrow \infty} D \xi_n = 0$, то

$$\lim_{n \rightarrow \infty} \mathbb{E} (\xi_n - C)^2 = \lim_{n \rightarrow \infty} \mathbb{E} (\xi_n - \mathbb{E} \xi_n)^2 = \lim_{n \rightarrow \infty} D \xi_n = 0. \quad \blacktriangleright$$

§5 Закон великих чисел

Ймовірність є ключовим інструментом для аналізу невизначеності, що дозволяє прогнозувати результати подій і приймати обґрунтовані рішення в багатьох сферах – від промислового виробництва до фінансового планування та управління ризиками. Події, ймовірність настання яких близька до одиниці, вважаються майже гарантованими, що забезпечує надійну основу для планування. Наприклад, у виробничих процесах знання про те, що певний продукт із ймовірністю 0,99 відповідатиме стандартам якості, дає змогу планувати ресурси та мінімізувати втрати. Аналогічно, в управлінні ризиками фінансовий аналітик може покладатися на високу ймовірність стабільності певного активу, щоб розробити ефективну інвестиційну стратегію. Така передбачуваність є ключовою для прийняття обґрунтованих рішень у ситуаціях, де невизначеність відіграє значну роль.

Однак визначити, яка саме ймовірність достатня, щоб вважати подію „достовірною”, неможливо однозначно. Усе залежить від контексту та практичних вимог. Наприклад, у медичних дослідженнях імовірність 0,95 може бути достатньою для затвердження ліків, тоді як у ядерній енергетиці вимоги до безпеки можуть вимагати ймовірностей, ближчих до 0,9999.

Події з малою, але додатною ймовірністю також відіграють важливу роль. Хоча ймовірність одиначної такої події (наприклад, збій обладнання) може бути мізерною, при великій кількості випробувань чи спостережень імовірність її настання хоча б раз зростає. Наприклад, якщо ймовірність поломки машини становить 0,01 за місяць, то за кілька років ризик хоча б однієї поломки стає значним, що вимагає ретельного планування.

Ця двоїста природа ймовірностей – висока передбачуваність подій з імовірністю, близькою до одиниці, та потенційна значущість

рідкісних подій – робить теорію ймовірностей незамінною для розуміння поведінки випадкових процесів. Одним із центральних результатів цієї теорії є закон великих чисел, який описує, як поведуться середні значення випадкових величин при великій кількості спостережень. Закон великих чисел стверджує, що середнє значення результатів великої кількості незалежних випробувань із високою ймовірністю буде близьким до математичного сподівання відповідної випадкової величини. Це означає, що випадкові відхилення, які можуть бути суттєвими в невеликих вибірках, наприклад, у перших кількох кидках монети, поступово згладжуються при збільшенні кількості спостережень. У результаті ми отримуємо стабільні та передбачувані значення, які можна використовувати для прогнозування.

Практичне значення закону великих чисел важко переоцінити. У статистиці він дозволяє робити висновки про характеристики популяції на основі вибірових даних, зменшуючи вплив випадкових відхилень, зокрема, оцінювати середній зріст людей у країні за результатами опитування. У страхуванні закон великих чисел допомагає компаніям точно прогнозувати середні виплати, аналізуючи дані про велику кількість клієнтів. Навіть у повсякденному житті цей принцип проявляється, як-от казино, де стабільна перевага закладу забезпечується великою кількістю ігор, попри випадкові виграші гравців.

Закон великих чисел існує у двох основних формах: слабкій і сильній. Слабкий закон великих чисел стверджує, що ймовірність значного відхилення середнього значення вибірки від математичного сподівання прямує до нуля при зростанні кількості випробувань. Сильний закон великих чисел іде далі, гарантуючи, що середнє значення вибірки майже напевно (тобто з ймовірністю 1) збігається до математичного сподівання при необмеженому збільшенні кількості спостережень. Ці формулювання роблять закон великих чисел важливим інструментом для аналізу довгострокової поведінки випадкових процесів.

Отже, закон великих чисел є фундаментальним принципом теорії ймовірностей, який забезпечує передбачуваність і стабільність у світі невизначеності. Його застосування охоплює широкий спектр дисциплін – від науки і техніки до економіки та соціальних наук, де він допомагає будувати надійні моделі, оптимізувати ресурси та зменшувати ризики. Завдяки цьому закону ми можемо з впевненістю аналізувати довгострокову поведінку випадкових процесів і приймати обґрунтовані рішення на основі великих обсягів даних.

Це загальне уявлення про теорему типу закону великих чисел можна сформулювати і дещо конкретніше. Нехай задана деяка послідовність випадкових величин $\{\xi_n\}_1^\infty$. Вважається, що існують математичні сподівання $\mathbb{E}\xi_n$.

Законом великих чисел називається всяке твердження, яке стосується поведінки при $n \rightarrow \infty$ випадкової величини

$$\eta_n = \frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n \mathbb{E} \xi_i.$$

Іноді мова йде навіть про сукупність тверджень. Кажуть – закон великих чисел, бо $n \rightarrow \infty$. Найпростішим із таких тверджень є відома теорема Чебишова (закон великих чисел у формі Чебишова).

Теорема Чебишова (закон великих чисел у формі Чебишова). Нехай випадкові величини ξ_n – незалежні, однаково розподілені, існують математичні сподівання $\mathbb{E}\xi_n$ (для кожного n), а також існує стала $C > 0$, така, що для всіх n $D\xi_n \leq C$. Тоді $\eta_n \xrightarrow[n \rightarrow \infty]{p} 0$.

◀ Оскільки випадкові величини ξ_n однаково розподілені, то їх математичні сподівання однакові. Позначимо $\mathbb{E}\xi_n = a$ (n довільне). Тоді

$$\eta_n = \frac{1}{n} \sum_{i=1}^n \xi_i - a.$$

Обчислимо (використовуючи властивості математичного сподівання і дисперсії) математичне сподівання і дисперсію η_n :

$$\mathbb{E} \eta_n = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \xi_i - a = a - a = 0,$$

$$D \eta_n = \frac{1}{n^2} \sum_{i=1}^n D \xi_i \leq \frac{1}{n^2} \sum_{i=1}^n C = \frac{C}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Звідси, за властивістю **6⁰** §4, одержується, що η_n прямує до нуля у середньому квадратичному, а тому і за ймовірністю. ►

Тобто встановлено, що

$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n \mathbb{E} \xi_i \right| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0. \quad (*)$$

У літературі доведено серію теорем, в яких не вимагається однакої розподіленості, а вимагається лише існування математичних сподівань $\mathbb{E} \xi_i$, а також теорем, в яких умова однакових розподілів величин ξ_i забезпечує виконання вписаного співвідношення без вимоги $D \xi_i \leq C$ (теорема Хінчина^{*}). Аналогічний результат доведено за умови $\frac{1}{n} \sum_{i=1}^n D \xi_i \xrightarrow{n \rightarrow \infty} 0$. Відомі також необхідні і достатні умови виконання співвідношення (*).

Із теорем, які складають закон великих чисел, випливає правило оцінювання (знаходження наближеного значення) математичного сподівання деякої випадкової величини ξ . Припустимо, що у результаті експерименту знайдено (визначено) $\xi_1^0, \xi_2^0, \dots, \xi_n^0$ – конкретні реалізації випадкової величини ξ . Причому експеримент проводився в однакових умовах і без систематичних похибок. Можемо вважати, що проводився один експеримент над кожною із випадкових величин

^{*} Хінчин Олександр Якович (1894-1959) – радянський математик.

$\xi_1, \xi_2, \dots, \xi_n$, які розподілені за тим самим законом, що й випадкова величина ξ . Якщо ці експерименти незалежні, то можемо вважати, що випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ також незалежні. Тобто правомірно стверджувати, що ξ_1^0 – це реалізація (конкретне значення в результаті експерименту) випадкової величини ξ_1 , ξ_2^0 – випадкової величини ξ_2 і т.д., ξ_i^0 – реалізація випадкової величини ξ_i . Але вони однаково розподілені, якщо спостереження позбавлені систематичних похибок, тобто $\mathbb{E}\xi = \mathbb{E}\xi_i = a$.

Тому, використовуючи співвідношення (*), можемо записати

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n \xi_i^0 - \mathbb{E}\xi\right| > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0,$$

а звідси

$$\mathbb{E}\xi \approx \frac{1}{n}\sum_{i=1}^n \xi_i^0$$

– це так зване **правило середнього арифметичного** у теорії обробки результатів: якщо проводити багато експериментів, то математичне сподівання приблизно дорівнює середньому арифметичному отриманих результатів експерименту.

Якщо проведення великої кількості експериментів неможливе або досить дороге, то позначимо

$$a_0 = \xi_1^0,$$

якщо проведено ще один експеримент, то

$$a_1 = \frac{\xi_1^0 + \xi_2^0}{2} = \frac{a_0 + \xi_2^0}{2},$$

далі

$$a_2 = \frac{\xi_1^0 + \xi_2^0 + \xi_3^0}{3} = \frac{2a_1 + \xi_3^0}{3}, \quad a_3 = \frac{\xi_1^0 + \xi_2^0 + \xi_3^0 + \xi_4^0}{4} = \frac{3a_2 + \xi_4^0}{4},$$

тому

$$a_n = \frac{\xi_1^0 + \dots + \xi_n^0}{n+1} = \frac{na_{n-1} + \xi_{n+1}^0}{n+1}.$$

Можемо вважати $(\mathbb{E}\xi)_n$, як n -е наближення для математичного сподівання, таким:

$$(\mathbb{E}\xi)_n = \frac{n(\mathbb{E}\xi)_{n-1} + \xi_{n+1}}{n+1}.$$

Тому, згідно із законом великих чисел, $p\lim_{n \rightarrow \infty} (\mathbb{E}\xi)_n = \mathbb{E}\xi$.

Закон великих чисел служить обґрунтуванням серії обчислювальних методів (методів Монте Карло). Наприклад, нехай необхідно обчислити $\int_a^b g(x)dx$, де $g(x)$ – деяка неперервна функція.

Можливо, що аналітичний вигляд $g(x)$ невідомий, а можемо знайти лише значення $g(x)$ у деяких точках (наприклад, у результаті експерименту).

Нехай випадкові величини ξ_n розподілені рівномірно на $[a, b]$, тобто

$$f_{\xi_n}(x) = \begin{cases} 0, & x \notin [a, b]; \\ \frac{1}{b-a}, & x \in [a, b]. \end{cases}$$

і незалежні.

Можна поставити питання про знаходження математичного сподівання випадкових величин $g(\xi_n)$:

$$\mathbb{E}g(\xi_n) = \int_a^b \frac{1}{b-a} g(x)dx = \frac{1}{b-a} \int_a^b g(x)dx.$$

Тому

$$\int_a^b g(x)dx = (b-a) \cdot \mathbb{E}g(\xi_n) \approx (b-a) \frac{1}{n} \sum_{i=1}^n g(\xi_i^0).$$

Це зручно для обчислювальних експериментів і досить просто. Описаний метод називається **методом статистичних випробувань** обчислення інтегралів (методом Монте-Карло).

З теореми Чебишова випливає (є наслідком) відома **теорема Бернуллі**. Дійсно, нехай проводиться серія випробувань, у кожному з яких певна подія відбувається з імовірністю p і не відбувається з імовірністю $q=1-p$, $p \geq 0$, $q \geq 0$. Для $i=1,2,\dots,n$ можемо розглянути випадкові величини

$$\xi_i = \begin{cases} 1, & \text{якщо є успіх в } i\text{-му експерименті (подія відбулася),} \\ 0, & \text{якщо неуспіх в } i\text{-му експерименті (подія не відбулася).} \end{cases}$$

Тобто закон розподілу випадкових величин ξ_i такий:

ξ_i	0	1
P	q	p

Тоді $\mathbb{E}\xi_i = p$, $D\xi_i = pq$ для всіх $i=1,2,\dots,n$.

Розглянемо випадкову величину $\frac{1}{n} \sum_{i=1}^n \xi_i = \nu_n$, яка є частотою успіхів у перших n випробуваннях (відношенням кількості появ події до кількості випробувань). Оскільки $\mathbb{E}\xi_i = p$, то $\frac{1}{n} \sum_{i=1}^n \mathbb{E}\xi_i = p$. Отже, для послідовності $\{\xi_n\}$ виконані всі умови теореми Чебишова, тому для будь-якого $\varepsilon > 0$

$$P\{|\nu_n - p| > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0,$$

тобто частота події за ймовірністю прямує до ймовірності події:

$$p \lim_{n \rightarrow \infty} \nu_n = p.$$

Зміст цього твердження полягає у тому, що введене означення ймовірності відповідає інтуїтивному розумінню ймовірності як границі частоти.

Із теорем закону великих чисел випливають відомі теореми Муавра-Лапласа і Пуассона.

Сформульовані теореми встановлюють факт і умови збіжності за ймовірністю тих чи інших випадкових величин до постійних, не випадкових величин.

Закон великих чисел відіграє важливу роль у практичних застосуваннях теорії ймовірностей. Властивість випадкових величин при певних умовах поводити себе практично як не випадкові дозволяє впевнено оперувати з цими величинами, передбачати результати масових випадкових явищ.

Можливість таких передбачень у масових випадкових явищах ще більше розширюється наявністю іншої групи граничних теорем, які стосуються вже не граничних значень випадкових величин, а граничних законів розподілів.

§6. Центральна гранична теорема

У цьому параграфі розглядається одна з групи теорем, відомих під назвою „центральної граничної теореми”, яку інколи називають „кількісною формою закону великих чисел”.

Всі форми центральної граничної теореми присвячені встановленню умов, при яких виникає нормальний закон розподілу. Оскільки ці умови на практиці досить часто виконуються, нормальний закон розподілу є найбільш розповсюдженим із законів розподілу, який найчастіше зустрічається у випадкових явищах природи. Він виникає у всіх випадках, коли випадкова величина, яку досліджують, може бути представлена у вигляді суми досить великої кількості незалежних (або слабко залежних) елементарних складових, кожна з яких окремо порівняно мало впливає на суму.

Центральна гранична теорема є одним із фундаментальних принципів теорії ймовірностей і математичної статистики, який відіграє ключову роль у розумінні поведінки вибірових середніх. Вона стверджує, що сума великої кількості незалежних і однаково розподілених випадкових величин, незалежно від їхнього початкового розподілу, за певних умов наближається до нормального розподілу.

Центральна гранична теорема пояснює, чому нормальний розподіл настільки поширений у природі та статистичних дослідженнях, і слугує основою для багатьох статистичних методів, таких як побудова довірчих інтервалів чи перевірка гіпотез. Ця теорема не лише має теоретичне значення, але й широко застосовується в економіці, природничих науках, соціології та інших дисциплінах, де аналізуються великі обсяги даних. Її практична цінність полягає в можливості спрощення складних статистичних обчислень і прогнозування на основі вибірових даних.

Теорема (Ляпунова. Центральна гранична). Нехай $\{\xi_i\}$ – послідовність незалежних однаково розподілених випадкових величин, таких, що $\mathbb{E}\xi_i = 0$, $D\xi_i = 1$. Нехай

$$\eta_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i,$$

і $F_{\eta_n}(x)$ – функція розподілу випадкової величини η_n . Границя $F_{\eta_n}(x)$ при $n \rightarrow \infty$ існує і дорівнює $\lim_{n \rightarrow \infty} F_{\eta_n}(x) = F(x)$ (збіжність розуміють в основному, тобто у точках неперервності), де $F(x)$ – функція нормального розподілу з параметрами 0 і 1:

$$\lim_{n \rightarrow \infty} F_{\eta_n}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Коментар: якщо багато факторів впливають на деякий процес, то сумарний (результуючий) вплив – це нормально розподілена випадкова величина.

◀ Згідно з властивостями характеристичних функцій, досить довести, що послідовність характеристичних функцій для випадкової величини η_n має границею $e^{-\frac{t^2}{2}}$, тобто $\lim_{n \rightarrow \infty} \varphi_{\eta_n}(t) = e^{-\frac{t^2}{2}}$.

Позначимо $\varphi_i(t) = \varphi_{\xi_i}(t) = \varphi(t)$ (індекс i можна зняти, бо ξ_i однаково розподілені і всі характеристичні функції $\varphi_i(t)$ однакові). За властивостями характеристичних функцій

$$\varphi_{\eta_n}(t) = \left[\varphi\left(\frac{t}{\sqrt{n}}\right) \right]^n,$$

а також

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \varphi''(0)\frac{t^2}{2} + \frac{t^2}{2}\varepsilon(t), \quad \varepsilon(t) \xrightarrow{t \rightarrow 0} 0.$$

Але за властивостями характеристичних функцій,

$$\varphi(0) = 1,$$

$$\varphi'(0) = i \cdot \mathbb{E} \xi_i = 0,$$

$$\varphi''(0) = i^2 \cdot \mathbb{E} \xi_i^2 = i^2 (D\xi_i + (\mathbb{E} \xi_i)^2) = i^2 D\xi_i = -1.$$

Отже,

$$\varphi(t) = 1 - \frac{t^2}{2} + \frac{t^2}{2} \varepsilon(t) = 1 - \frac{t^2}{2} (1 - \varepsilon(t)).$$

Далі маємо таке

$$\varphi_{\eta_n}(t) = \left[1 - \frac{t^2}{2n} \left(1 - \varepsilon\left(\frac{t}{\sqrt{n}}\right) \right) \right]^n.$$

Для будь-якого фіксованого t при $n \rightarrow \infty$ очевидно, що $\varepsilon\left(\frac{t}{\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} 0$.

Тому

$$\lim_{n \rightarrow \infty} \varphi_{\eta_n}(t) = \lim_{n \rightarrow \infty} \left[1 - \frac{t^2}{2n} \right]^n = e^{-\frac{t^2}{2}},$$

а це – характеристична функція випадкової величини нормально розподіленої з параметрами 0 і 1. ►

Якщо б випадкові величини ξ_i не були б стандартизовані, то можна було б перейти до стандартизованих випадкових величин

$$\bar{\xi}_i = \frac{\xi_i - \mathbb{E} \xi_i}{\sigma_{\xi_i}} = \frac{\xi_i - a}{\sigma}$$

і тоді

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\sigma \bar{\xi}_i + a) = \sigma \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{\xi}_i \right) + a\sqrt{n} = \sigma \eta_n + a\sqrt{n},$$

η_n – в границі розподілена нормально з параметрами $N(0, 1)$. Тому в границі сума буде розподілена нормально з параметрами $N(a\sqrt{n}, \sigma^2)$.

У більш загальному випадку, коли випадкові величини незалежні, але не обов'язково однаково розподілені, в умовах центральної граничної теореми для того, щоб нормована сума була розподілена приблизно як нормальний розподіл при великому n , вимагається наявність моменту вище другого порядку та збіжність до нуля нормованої суми.

Можна отримати необхідні і достатні умови того, що сума випадкових величин в границі при $n \rightarrow \infty$ розподілена нормально.

Необхідними і достатніми умовами виконання центральної граничної теореми є **умови Ліндеберга***:

для будь-якого $\tau > 0$ границя

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-m_k| > \tau B_n} (x-m_k)^2 f_k(x) dx = 0,$$

де $m_k = \mathbb{E}\xi_k$ – математичне сподівання ξ_k , $f_k(x) = f_{\xi_k}(x)$ – щільність

розподілу ξ_k , $B_n = \sqrt{\sum_{k=1}^n D\xi_k}$, $D\xi_k$ – дисперсія ξ_k .

* Ярл Вальдемар Ліндеберг (Jarl Waldemar Lindeberg, 1876-1932) – фінський математик.

РОЗДІЛ 4

МАТЕМАТИЧНА СТАТИСТИКА

§ 1. Предмет і задачі математичної статистики

Статистика – це наука, яка займається реєстрацією, описом і аналізом даних, які отримуються в результаті спостереження за масовими випадковими явищами.

Умовно з цього означення можна виділити три напрями статистики:

- перший напрям (завдання, задача статистики) – збір даних, які характеризують явище;
- другий напрям – дослідження одержаних даних з метою встановлення певних закономірностей;
- третій напрям – розробка прийомів спостереження, методів збору даних і методів аналізу результатів.

Остання задача – це суть математичної статистики.

Теоретичною основою (базою) математичної статистики є теорія ймовірностей. На основі результатів спостережень випадкової події у математичній статистиці робляться певні висновки про розподіл випадкової величини (події) або про деякі параметри, що характеризують розподіл (від події до величини можна перейти).

Нехай спостереження ведуться за деякою випадковою величиною $\xi(\omega)$ (наприклад, силою струму, швидкістю вітру і т.д.). Множина всіх можливих значень $\xi(\omega)$ (ω – деякий неконтрольований фактор) називається **генеральною сукупністю** такої випадкової величини.

Вважаємо, що спостереження незалежні. Нехай в результаті n спостережень одержано n значень $\xi(\omega)$: $x_1, x_2, \dots, x_n$ (серед них можуть бути й однакові).

Набір чисел $x_1, x_2, \dots, x_n$ називається **вибіркою** із генеральної сукупності, число n – **об'ємом вибірки**. Якщо здійснювати іншу серію спостережень (експериментів), то набір чисел може змінитися, тобто вибірка стане іншою. Тому природно вважати, що числа x_i , $i = \overline{1, n}$ є реалізаціями випадкових величин $\xi_i(\omega)$, розподіли яких збігаються з розподілом $\xi(\omega)$. Як правило, спостереження і їх результати незалежні. Тому випадкові величини $\xi_i(\omega)$ слід вважати незалежними. Отже, практично, вибірка – це реалізація векторної випадкової величини $(\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega))$, компоненти якої незалежні і розподілені так, як $\xi(\omega)$.

Формально **завдання** (ціль, мета) **математичної статистики** полягає в:

1) знаходженні (оцінюванні) розподілу (закону розподілу) випадкової величини $\xi(\omega)$ за реалізацією вектора

$$(\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega));$$

2) оцінюванні (обчисленні) параметрів розподілу випадкової величини $\xi(\omega)$;

3) оцінюванні (знаходженні) числових характеристик $\xi(\omega)$;

4) якщо вивчаються дві випадкові величини (два явища) $\xi(\omega)$ і $\eta(\omega)$, то необхідно знайти або якось оцінити ступінь залежності між цими випадковими величинами (залежні чи ні, наскільки тісно залежні). Може виникати питання про розподіл $\xi(\omega)$ і $\eta(\omega)$;

5) якщо з деяких міркувань висловлюється гіпотеза про розподіл випадкової величини $\xi(\omega)$, то необхідно перевірити, чи результати спостережень не заперечують цю гіпотезу (мова не йде про підтвердження гіпотези, а лише про те, відкинути чи не відкинути цю гіпотезу).

Розглянемо спочатку задачу 2).

Нехай функція розподілу випадкової величини ξ залежить від деякого параметра $\theta: F_{\xi}(x, \theta)$. Вигляд цієї функції відомий, невідомий лише параметр. Необхідно за вибіркою випадкової величини ξ знайти параметр θ . Тобто треба знайти якесь правило (функцію) $\theta = h_n(x_1, x_2, \dots, x_n)$, за яким можна отримати θ , знаючи $x_1, x_2, \dots, x_n$. Взагалі це зробити навряд чи вдасться, бо при інших експериментах одержимо інші значення x_i . Будемо робити так: знаходити число (випадкову величину) $\hat{\theta} = h_n(x_1, x_2, \dots, x_n)$. По суті $\hat{\theta}$ є випадковою величиною, тобто треба писати $\hat{\theta} = \hat{\theta}_n(\omega)$. Будемо ставити питання про те, наскільки значення випадкової величини $\hat{\theta}$ близьке до значення θ .

Функція $h_n(x_1, x_2, \dots, x_n)$ називається **статистикою** для параметра θ , тобто можна покласти $h_n(x_1, x_2, \dots, x_n) = \hat{\theta}$. **На статистики накладаються такі вимоги:**

1. Статистика повинна бути незміщеною (незсунутою), тобто

$$\mathbb{E} \hat{\theta}_n(\omega) = \mathbb{E} \hat{\theta} = \theta.$$

1'. Іноді виставляється слабша умова: статистики повинні бути асимптотично незміщені, тобто

$$\lim_{n \rightarrow \infty} \mathbb{E} \hat{\theta} = \theta.$$

2. Статистики повинні бути суттєвими (істотними) **конзистентними**. Це означає, що границя за ймовірністю дорівнює θ :

$$p \lim_{n \rightarrow \infty} \hat{\theta}_n(\omega) = \theta.$$

3. Статистики повинні бути ефективними. Це означає, що дисперсія $\hat{\theta}_n(\omega)$ мінімальна: $D \hat{\theta}_n(\omega)$ – є мінімальною. (Якщо пропонувати різні статистики, то вони повинні задовольняти 1–2 і серед них треба вибирати лише ті, для яких дисперсія мінімальна).

Приклад (на побудову статистики). Нехай відомо, що деяка скалярна випадкова величина $\xi(\omega)$ розподілена рівномірно. Невідомий відрізок $[a, b]$. Відома щільність (при $a < b$):

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

Параметри a і b невідомі. Треба їх знайти (оцінити) за спостереженнями над випадковою величиною $\xi(\omega)$. Нехай проведено n спостережень і отримано вибірку $x_1, x_2, \dots, x_n$. Запропонуємо деякі статистики:

$$\hat{\theta}_1 = \max_{1 \leq i \leq n} x_i, \quad \hat{\theta}_2 = \min_{1 \leq i \leq n} x_i,$$

$$\hat{\theta}_3 = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\theta}_4 = \frac{1}{2}(x_1 + x_n).$$

Можна запропонувати ще й інші статистики. Поставимо за мету використати ці статистики, щоб знайти b – за статистикою $\hat{\theta}_1$, a – за $\hat{\theta}_2$, математичне сподівання $\xi(\omega)$ – за $\hat{\theta}_3$.

Перевіримо незміщеність запропонованих статистик. Потрібно знайти математичне сподівання. Випишемо функцію розподілу $\hat{\theta}_1$:

$$\begin{aligned} F_{\hat{\theta}_1}(x) &= P\{\hat{\theta}_1 < x\} = P\left\{\max_{1 \leq i \leq n} x_i < x\right\} = \\ &= P\{x_1 < x, x_2 < x, \dots, x_n < x\} = \end{aligned}$$

де x_i – це числа, а як реалізації – це випадкові величини, причому незалежні, тому

$$= \prod_{i=1}^n P\{x_i < x\} = \prod_{i=1}^n \int_{-\infty}^x f_{\xi}(y, a, b) dy = \left(\int_{-\infty}^x f_{\xi}(y, a, b) dy \right)^n.$$

Щільність розподілу $\hat{\theta}_1$ тоді така:

$$\begin{aligned}
 f_{\hat{\theta}_1}(x, a, b) &= n \left(\int_{-\infty}^x f_{\xi}(y, a, b) dy \right)^{n-1} \cdot f_{\xi}(x, a, b) = \\
 &= \begin{cases} 0, & x \notin [a, b], \\ n \left(\int_a^x \frac{1}{b-a} dy \right)^{n-1} \cdot \frac{1}{b-a}, & x \in [a, b] \end{cases} = \\
 &= \begin{cases} 0, & x \notin [a, b], \\ \frac{n(x-a)^{n-1}}{(b-a)^n}, & x \in [a, b]. \end{cases}
 \end{aligned}$$

Аналогічно можна встановити, що

$$f_{\hat{\theta}_2}(x, a, b) = \begin{cases} 0, & x \notin [a, b], \\ \frac{n(b-x)^{n-1}}{(b-a)^n}, & x \in [a, b]. \end{cases}$$

Тоді математичне сподівання $\hat{\theta}_1$ дорівнює:

$$\begin{aligned}
 \mathbb{E} \hat{\theta}_1 &= \int_a^b x \frac{n(x-a)^{n-1}}{(b-a)^n} dx = n \int_a^b \frac{(x-a)^n}{(b-a)^n} dx + n \int_a^b \frac{a(x-a)^{n-1}}{(b-a)^n} dx = \\
 &= \frac{n(b-a)^{n+1}}{(n+1)(b-a)^n} + a = \frac{n}{n+1}(b-a) + \frac{n+1}{n+1}a = \frac{n}{n+1}b + \frac{1}{n+1}a.
 \end{aligned}$$

Якщо статистикою $\hat{\theta}_1$ оцінювати правий кінець b відрізка, то оцінка зміщена, але асимптотично незміщена оцінка параметра b :

$$\mathbb{E} \hat{\theta}_1 \neq b, \text{ але } \lim_{n \rightarrow \infty} \mathbb{E} \hat{\theta}_1 = b.$$

Аналогічно, якщо статистикою $\hat{\theta}_2$ оцінювати лівий кінець a відрізка, то знову оцінка зміщена, але асимптотично незміщена:

$$\mathbb{E} \hat{\theta}_2 \neq a, \text{ але } \lim_{n \rightarrow \infty} \mathbb{E} \hat{\theta}_2 = a.$$

Подібним способом можна показати, що ці оцінки асимптотично суттєві.

Розглянемо третю оцінку $\hat{\theta}_3$. Її математичне сподівання дорівнює

$$\mathbb{E} \hat{\theta}_3 = \frac{1}{n} \sum_{i=1}^n \mathbb{E} x_i = \frac{1}{n} \sum_{i=1}^n \frac{a+b}{2} = \frac{a+b}{2}.$$

Отже, статистику $\hat{\theta}_3$, з точки зору незміщеності, можна використати як оцінку математичного сподівання (ця оцінка незміщена). Чи є вона істотна? Для результатів спостережень x_i існують математичні сподівання, дисперсії обмежені і дорівнюють $\frac{(a+b)^2}{12}$ (бо вони також рівномірно розподілені на $[a, b]$). Згідно із законом великих чисел,

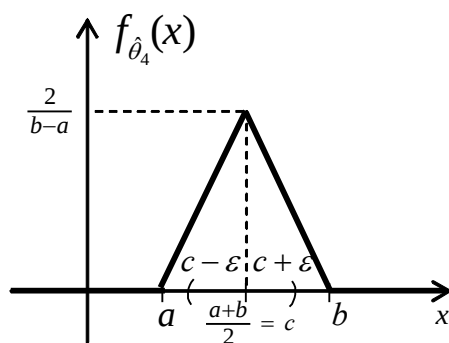
$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n} \sum_{i=1}^n \mathbb{E} x_i \right| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0.$$

Тобто $\hat{\theta}_3$ є істотною оцінкою для математичного сподівання.

Перейдемо до оцінки $\hat{\theta}_4$ і запишемо її щільність розподілу:

$$f_{\hat{\theta}_4}(x, a, b) = \begin{cases} 0, & x < a, \\ \frac{4(x-a)}{(b-a)^2}, & a \leq x \leq \frac{a+b}{2}, \\ \frac{4(b-x)}{(b-a)^2}, & \frac{a+b}{2} < x \leq b, \\ 0, & x > b. \end{cases}$$

Графік щільності розподілу такий:



Крім цього, математичне сподівання $\mathbb{E} \hat{\theta}_4 = \frac{a+b}{2}$, тобто оцінка $\hat{\theta}_4$ – незміщена. Перевіримо, чи оцінка $\hat{\theta}_4$ суттєва.

$$P\left\{\left|\hat{\theta}_4 - \frac{a+b}{2}\right| > \varepsilon\right\} = \int_{x: \left|x - \frac{a+b}{2}\right| > \varepsilon} f_{\hat{\theta}_4}(x) dx \xrightarrow{n \rightarrow \infty} 0,$$

тобто $\hat{\theta}_4$ не є конзистентною, асимптотично суттєвою. У прикладі, що розглядається, як оцінку, яка задовольняє умови статистик, можна вибрати тільки одну статистику $\hat{\theta}_3$

$$\mathbb{E} \xi = \frac{a+b}{2} \approx \frac{1}{n} \sum_{i=1}^n x_i.$$

Тому для математичного сподівання випадкової величини $m = \mathbb{E} \xi$ можна використати таку оцінку (не залежно від розподілу випадкової величини ξ):

$$m = \mathbb{E} \xi \approx m^* = h^*(x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i.$$

Якщо для невідомого параметра θ вибрана деяка статистика (оцінка), то природно з'ясувати точність цієї оцінки. Як міру точності можна використовувати

$$\mathbb{E}(\hat{\theta} - \theta)^2,$$

де $\hat{\theta}$ – деяка вибрана статистика, θ – точне значення параметра. Якщо $\hat{\theta}$ незміщена, тобто $\mathbb{E}\hat{\theta} = \theta$, то ця міра точності $\mathbb{E}(\hat{\theta} - \theta)^2$ є дисперсією $\hat{\theta}$: $\mathbb{E}(\hat{\theta} - \theta)^2 = D\hat{\theta}$. Тобто оцінка точності мінімальна (математичне сподівання мінімальне), коли дисперсія мінімальна.

Припустимо, що розглядається множина всіх можливих незміщених оцінок для деякого параметра θ . Позначимо цю множину

$$D_{\theta} = \{h_{\theta}(x_1, \dots, x_n)\}.$$

Оскільки кожна статистика незміщена, то для кожного елемента цієї множини оцінкою точності є дисперсія цієї статистики.

Деяка статистика (оцінка) θ^* називається **ефективною**, якщо дисперсія

$$D\theta^* = \inf_{\hat{\theta} \in D_{\theta}} D\hat{\theta}.$$

Тобто ефективною називається та оцінка, на якій досягається нижня грань дисперсій.

Якщо б θ^* була відома, то кожному з конкретних оцінок можна було б порівнювати з θ^* . Завжди буде $D\hat{\theta} \geq D\theta^*$. Практично дуже важко знайти θ^* , і навіть $D\theta^*$.

Зрозуміло, що за скінченною кількістю значень $x_1, x_2, \dots, x_n$, отриманих у результаті експерименту, взагалі кажучи, неможливо повною мірою вивчити випадкову величину. Як за допомогою цієї вибірки $x_1, x_2, \dots, x_n$ оцінити випадкову величину? Певною мірою відповідь на питання про можливість „якісного” оцінювання невідомого параметра дає нерівність Крамера* -Рао**.

* Карл Харальд Крамер (Karl Harald Cramér, 1893-1985) – шведський математик і статистик.

** Кальямпуді Радхакрішна Рао (Calyampudi Radhakrishna Rao, нар.1920) – індійський математик і статистик.

Вважається, що щільність сумісного розподілу $f(x_1, x_2, \dots, x_n, \theta)$ досить регулярно залежить від параметра θ (на практиці щільність сумісного розподілу не відома, її треба будувати).

Позначимо: вектор $\bar{x} = (x_1, x_2, \dots, x_n)$, h – статистика, яка оцінюється.

Теорема (Крамера-Рао). Якщо існують похідні

$$\frac{\partial}{\partial \theta} f(\bar{x}, \theta), \frac{\partial^2}{\partial \theta^2} f(\bar{x}, \theta)$$

та інтеграли

$$\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial \theta} f(\bar{x}, \theta) \right| d\bar{x} < \infty, \quad \int_{\mathbb{R}^n} \left| \frac{\partial^2}{\partial \theta^2} f(\bar{x}, \theta) \right| d\bar{x} < \infty,$$

$$\int_{\mathbb{R}^n} \left| h(\bar{x}) \frac{\partial}{\partial \theta} f(\bar{x}, \theta) \right| d\bar{x} < \infty$$

і математичні сподівання

$$\mathbb{E} \left| \frac{\partial}{\partial \theta} \ln f(\bar{x}, \theta) \right|^2 < \infty, \quad \mathbb{E} \left| \frac{\partial^2}{\partial \theta^2} \ln f(\bar{x}, \theta) \right| < \infty,$$

то

$$D\theta^* \geq \frac{1}{\mathbb{E} \left(\frac{\partial}{\partial \theta} \ln f(\bar{x}, \theta) \right)^2}.$$

Виявляється, що дисперсія найкращої оцінки може бути оцінена знизу. Тобто кращої точності досягти не можна.

Знаменник у нерівності Крамера-Рао називається **кількістю інформації за Фішером***. Можна виписати умови, при яких така оцінка досягається.

* Сер Рональд Ейлер Фішер (Sir Ronald Aylmar Fisher, 1890-1962) – англійський статистик, біолог-еволюціоніст, генетик.

Якщо випробування незалежні, а випадкова величина ξ має щільність розподілу $g_{\xi}(x, \theta)$, то

$$f(\bar{x}, \theta) = \prod_{i=1}^n g_{\xi}(x_i, \theta)$$

і тоді нерівність Крамера-Рао набуває вигляду:

$$D\theta^* \geq \frac{1}{n \cdot \mathbb{E} \left(\frac{\partial}{\partial \theta} \ln g_{\xi}(x, \theta) \right)^2}.$$

Для цієї оцінки порядок малості не більший, ніж $\frac{1}{n}$.

Можна довести, що у випадку нормального розподілу порядок малості дорівнює $\frac{1}{n}$.

§2. Первинна обробка статистичної інформації

Нехай у результаті спостережень (експериментів) отримано вибірку $x_1, x_2, \dots, x_n$ з генеральної сукупності глибиною n . Серед x_i можуть бути однакові. За такою вибіркою можна побудувати так звану експериментальну функцію розподілу. Якщо

$$x_{\min} = \min\{x_1, \dots, x_n\}, \quad x_{\max} = \max\{x_1, \dots, x_n\},$$

то ця функція визначається так

$$\hat{F}_{\xi}(x) = \begin{cases} 0, & x \leq x_{\min}, \\ \sum_{i: x_i \leq x} \frac{\nu(x_i)}{n}, & x_{\min} < x \leq x_{\max}, \\ 1, & x > x_{\max}, \end{cases}$$

де $\nu(x_i)$ – кількість спостережень (експериментів), в яких виявилось значення x_i .

За цією формулою працювати не зручно. Потрібно провести первинну обробку вибірки. Першим кроком цієї обробки є побудова простого варіаційного ряду. Він має вигляд:

x'_1	x'_2	...	x'_k
$\nu(x'_1)$	$\nu(x'_2)$	...	$\nu(x'_k)$

де x'_i – різні значення отримані в результаті спостережень (різних значень є k , $k \leq n$, всі $x'_i \in \{x_1, \dots, x_n\}$). Взагалі кажучи, краще користуватися таким простим варіаційним рядом:

x'_1	x'_2	...	x'_k
p'_1	p'_2	...	p'_k

де $p'_i = \frac{\nu(x'_i)}{n}$ – частота появи значення x'_i . Ясно, що $\sum_{i=1}^k p'_i = 1$.

Надалі будемо вважати, що від самого початку задано простий варіаційний ряд і штрихи писати не будемо. Тобто об'єм вибірки n , у ній k різних значень $x_1 < x_2 < \dots < x_k$, впорядкованих за зростанням, частоти $p_i^* = \frac{\nu(x_i)}{n}$ і простий варіаційний ряд має вигляд:

x_1	x_2	...	x_k
p_1^*	p_2^*	...	p_k^*

Тоді експериментальна функція розподілу така:

$$\hat{F}_\xi(x) = \begin{cases} 0, & x \leq x_1, \\ \sum_{i: x_i < x} p_i^*, & x_1 < x \leq x_k, \\ 1, & x > x_k. \end{cases}$$

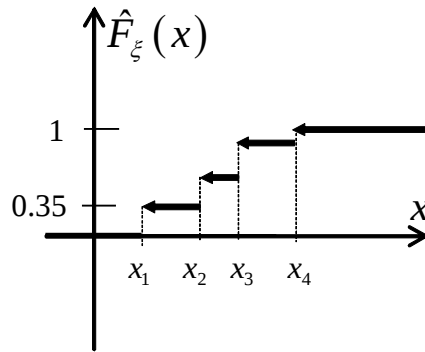
Приклад. Для простого варіаційного ряду

x_1	x_2	x_3	x_4
0,35	0,20	0,35	0,10

експериментальна функція розподілу записується так:

$$\hat{F}_\xi(x) = \begin{cases} 0, & x \leq x_1, \\ 0,35, & x_1 < x \leq x_2, \\ 0,55, & x_2 < x \leq x_3, \\ 0,90, & x_3 < x \leq x_4, \\ 1, & x > x_4. \end{cases}$$

і графік функції розподілу має такий вигляд:



Наступним кроком обробки інформації (особливо при великих n) є **побудова інтервального варіаційного ряду**. Це роблять так: відрізок $[x_{\min}, x_{\max}]$ ділять точками $x_{\min} = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_l = x_{\max}$ і будують таку таблицю

$[\alpha_0, \alpha_1)$	$[\alpha_1, \alpha_2)$	$\dots$	$[\alpha_{l-1}, \alpha_l]$
p_1^*	p_2^*	$\dots$	p_l^*

де p_i^* , $i = \overline{1, (l-1)}$ – частота попадання результатів спостережень на проміжок $[\alpha_{i-1}, \alpha_i)$, а p_l^* – частота попадання на відрізок $[\alpha_{l-1}, \alpha_l]$.

Це не єдина можливість. Розглядаються й інші варіанти побудови такого ряду:

- можна не включати у проміжок лівий кінець;
- для певного згладжування, при підрахунках частот p_i^* у випадках, коли $x_j = \alpha_j$, половину кількості появ x_j додають до лівого проміжку і половину кількості – до правого.

Позначимо $x_i^* = \frac{1}{2}(\alpha_{i-1} + \alpha_i)$, $i = \overline{1, l}$. У результаті отримується варіаційний ряд такого типу:

x_1^*	x_2^*	$\dots$	x_l^*
p_1^*	p_2^*	$\dots$	p_l^*

§3. Оцінювання числових характеристик

Як оцінку математичного сподівання випадкової величини ξ будемо використовувати m^* : якщо задана вибірка, то можна прийняти

$$m^* = \frac{1}{n} \sum_{i=1}^n x_i.$$

Якщо побудований простий варіаційний ряд, то

$$m^* = \sum_{i=1}^k x_i p_i^*.$$

Якщо побудований інтервальний варіаційний ряд, то як m^* можна використовувати ряд

$$m^* = \sum_{i=1}^l x_i^* p_i^*.$$

Ця оцінка незміщена й суттєва, як було показано в §1.

Як оцінку для дисперсії природно використати ту ж структуру – формально замінити знак математичного сподівання у виразі $D\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2$ на структуру схожу з m^* .

Наприклад, як D^* (для першого випадку) можна запропонувати таку статистику:

$$D^* = \frac{1}{n} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2.$$

Можна показати, що оцінка D^* дорівнює:

$$D^* = \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^n x_i \right)^2.$$

Знайдемо математичне сподівання запропонованої оцінки D^* . Бажано, щоб воно дорівнювало D , тобто щоб $\mathbb{E}D^* = D = D\xi$. Тоді ця оцінка виявилася б незміщеною.

$$\begin{aligned}\mathbb{E}D^* &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n x_i^2 - \left(\frac{1}{n}\sum_{i=1}^n x_i\right)^2\right) = \frac{1}{n}\left(\mathbb{E}\sum_{i=1}^n x_i^2 - \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n x_i\right)^2\right) = \\ &= \frac{1}{n}\left(\mathbb{E}\sum_{i=1}^n x_i^2 - \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n x_i^2 + 2\sum_{i<j} x_i x_j\right)\right) =\end{aligned}$$

вважаємо, що результати спостережень незалежні: $\mathbb{E}(x_i x_j) = \mathbb{E}x_i \mathbb{E}x_j$,
тому

$$= \frac{n-1}{n^2}\sum_{i=1}^n \mathbb{E}x_i^2 - \frac{2}{n^2}\sum_{i<j} \mathbb{E}x_i \mathbb{E}x_j =$$

але $\mathbb{E}x_i = \mathbb{E}x_j$, бо вони однаково розподілені

$$\begin{aligned}&= \frac{n-1}{n^2}\sum_{i=1}^n (\mathbb{E}x_i^2 - m^2 + m^2) - \frac{2}{n^2}\sum_{i<j} m \cdot m = \\ &= \frac{n-1}{n^2}\sum_{i=1}^n (D + m^2) - \frac{2}{n^2}m^2 \frac{n(n-1)}{2} =\end{aligned}$$

останній множник дорівнює кількості доданків в сумі $\sum_{i<j}$, тому

$$= \frac{n-1}{n^2}D \cdot n + \frac{n-1}{n^2}m^2 n - \frac{2}{n^2}m^2 \frac{n(n-1)}{2} = \frac{n-1}{n}D.$$

Отже, отримали: $\mathbb{E}D^* = \frac{n-1}{n}D \neq D$. Це означає, що запропонована оцінка (статистика) зміщена. Зауважимо, що вона асимптотично

незміщена, бо $\lim_{n \rightarrow \infty} \mathbb{E}D^* = D$. Отже, ця оцінка має ваду, яку треба виправити, щоб отримати незміщену оцінку. Тому слід використувати оцінку не D^* , а підправлену оцінку $D^{*'} = \frac{n}{n-1} D^*$. Для неї буде $\mathbb{E}D^{*'} = D$ і оцінка ця незміщена. Надалі будемо позначати (вважати):

$$m^* = \frac{1}{n} \sum_{i=1}^n x_i, \quad D^* = \frac{1}{n-1} \sum_{i=1}^n (x_i - m^*)^2.$$

Ці числа називаються **точковими оцінками для математичного сподівання і дисперсії**. Вони суттєві й незміщені.

Якщо задано простий варіаційний ряд, то

$$m^* = \sum_{i=1}^k x_i p_i^*, \quad D^* = \frac{n}{n-1} \sum_{i=1}^k (x_i - m^*)^2 p_i^*.$$

Якщо ж користуємось інтервальним варіаційним рядом, то ці оцінки будуть записуватись так:

$$m^{*'} = \sum_{i=1}^l x_i^* p_i^*, \quad D^{*'} = \frac{n}{n-1} \sum_{i=1}^l (x_i^* - m^{*'})^2 p_i^*.$$

При $n \rightarrow \infty$ m^* і $m^{*'}$ еквівалентні, а також D^* і $D^{*'}$ еквівалентні.

Аналогічно здійснюється побудова оцінок інших параметрів розподілу (моментів вищих порядків, ексцесу, моди і т.д.). При побудові інтервального варіаційного ряду, як правило, використовують l з відрізка $10 \leq l \leq 30$ або $l \approx 20$.

§4. Довірчий інтервал

Нехай θ – деякий параметр, що характеризує розподіл випадкової величини ξ . І нехай за певною вибіркою $x_1, x_2, \dots, x_n$ (отриманою після експериментів) побудована деяка оцінка θ^* параметра θ . Це, взагалі кажучи, функція від значень $x_1, x_2, \dots, x_n$: $\theta^* = \theta^*(x_1, x_2, \dots, x_n)$. Як було сказано вище, θ^* – випадкова величина. Можна поставити таке питання: наскільки знайдене значення θ^* наближає невідоме значення θ (яка точність)?

Питання можна поставити і так: чи не можна визначити якісь числа θ_1^* і θ_2^* ($\theta_1^* < \theta_2^*$), такі, щоб невідомий параметр (характеристика) θ попадав між ними ($\theta_1^* < \theta < \theta_2^*$) з високою ймовірністю? Цей інтервал (θ_1^*, θ_2^*) випадковий. Мова йде про те, щоб невідоме число θ „захоплювалося” цим інтервалом. Як правило, вимагають, щоб цей випадковий інтервал містив число θ з імовірністю, не меншою ніж 0,9 (це означає „висока ймовірність”).

Побудова вказаних інтервалів здійснюється так. Зафіксуємо деяке число β : $0 < \beta < 1$. Як правило, вибирають $\beta = 0,9$ або $0,95$ або $0,99$. Це число називають **довірчою ймовірністю**. Після цього визначається таке $\varepsilon > 0$, щоб виконувалась нерівність $P\{|\theta^* - \theta| < \varepsilon\} \geq \beta$. Якщо число ε знайдено (як саме знайдено, це інше питання), а це означає, що $P\{\theta^* - \varepsilon < \theta < \theta^* + \varepsilon\} \geq \beta$, то фактично побудовано інтервал

$$(\theta^* - \varepsilon, \theta^* + \varepsilon),$$

в якому з імовірністю, не меншою ніж β знаходиться невідомий параметр. Числа $\theta^* - \varepsilon$ і $\theta^* + \varepsilon$ називаються **довірчими границями**. Довжина інтервалу дорівнює 2ε .

Оскільки при розширенні інтервалу ймовірність зростає, то природно, що шукають найменше ε , або знаходять ε з рівності

$$P\{|\theta^* - \theta| < \varepsilon\} = \beta.$$

Розглянемо прості випадки. В загальнотеоретичному плані задача розв'язується дуже складно, бо θ^* є випадковою величиною і побудувати щільність розподілу цієї випадкової величини складно. Більше того, параметр θ невідомий. Задачу можна розв'язати наближено, якщо накладати на випадкову величину деякі додаткові умови. Наприклад, якщо вважати, що випадкова величина розподілена нормально, то невідомі всього два параметри – математичне сподівання і дисперсія. Тому мова може йти про знаходження довірчих інтервалів саме цих параметрів.

Є такі три підходи:

- математичне сподівання невідоме, дисперсія відома. Знайти довірчий інтервал для невідомого математичного сподівання;
- математичне сподівання невідоме і дисперсія невідома. Треба побудувати довірчий інтервал як для математичного сподівання, так і для дисперсії;
- математичне сподівання відоме, дисперсія невідома. Знайти довірчий інтервал для невідомої дисперсії.

Оскільки результати спостережень незалежні і величини x_i однаково розподілені, то, згідно з граничними теоремами, суми x_i при великих n розподілені нормально. Поставимо задачу про **побудову довірчого інтервалу для математичного сподівання m , якщо відома дисперсія D** . У цьому випадку m^* треба знайти з рівності:

$$P\{ |m^* - m| < \varepsilon \} = \beta.$$

Тобто

$$P\left\{ \frac{-\varepsilon}{\sqrt{Dm^*}} < \frac{m^* - m}{\sqrt{Dm^*}} < \frac{\varepsilon}{\sqrt{Dm^*}} \right\} = \beta.$$

Тоді випадкова величина $\frac{m^* - m}{\sqrt{Dm^*}}$ розподілена нормально з параметрами $N(0, 1)$. Знайдемо Dm^* :

$$Dm^* = D\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sum_{i=1}^n Dx_i = \frac{1}{n^2} nD.$$

Тобто

$$Dm^* = \frac{D}{n}, \quad \sqrt{Dm^*} = \frac{\sqrt{D}}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}}.$$

Таким чином, рівняння для знаходження ε має вигляд

$$P\left\{-\frac{\varepsilon\sqrt{n}}{\sigma} < \frac{m^* - m}{\sigma} \sqrt{n} < \frac{\varepsilon\sqrt{n}}{\sigma}\right\} = \beta.$$

У лівій частині цієї рівності записана ймовірність попадання в інтервал нормально розподіленої випадкової величини з параметрами $N(0, 1)$. Отже,

$$\beta = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\varepsilon\sqrt{n}}{\sigma}}^{\frac{\varepsilon\sqrt{n}}{\sigma}} e^{-\frac{x^2}{2}} dx = 2\Phi\left(\frac{\sqrt{n}}{\sigma} \varepsilon\right),$$

де $\Phi(x)$ – функція Лапласа: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$.

Таким чином, для того, щоб знайти ε , потрібно розв'язати рівняння $\Phi\left(\frac{\sqrt{n}}{\sigma} \varepsilon\right) = \frac{\beta}{2}$. Його розв'язують при допомозі таблиць – знаходять $\text{арс}\Phi\left(\frac{\beta}{2}\right)$ за таблицями значень функції $\Phi(x)$, тут $\text{арс}\Phi\left(\frac{\beta}{2}\right)$ – аргумент x функції $\Phi(x)$, такий, що $\Phi(x) = \frac{\beta}{2}$, тоді $\frac{\sqrt{n}}{\sigma} \varepsilon = \text{арс}\Phi\left(\frac{\beta}{2}\right)$.

Звідси $\varepsilon = \frac{\sigma}{\sqrt{n}} \text{арс}\Phi\left(\frac{\beta}{2}\right)$.

Тоді, як правило, пишуть так:

$$\varepsilon = \sigma \cdot t_{\beta}, \quad \text{де} \quad t_{\beta} = \frac{1}{\sqrt{n}} \operatorname{arc}\Phi\left(\frac{\beta}{2}\right).$$

Отже, довірчий інтервал для m , якщо відомо σ , має вигляд:

$$(m^* - \sigma t_{\beta}, m^* + \sigma t_{\beta}).$$

Якщо ж σ – **невідоме**, то для побудови інтервалу можна використати оцінку σ^* , і тоді **довірчий інтервал** буде таким:

$$(m^* - \sigma^* t_{\beta}, m^* + \sigma^* t_{\beta}),$$

де t_{β} знаходять за таблицями (таблиця 4 у Додатку В) за заданими n і β . При цьому вважається, що інтервал побудований грубо.

Аналогічно будується довірчий інтервал для дисперсії. Лише при цьому не можемо вважати, що доданки у виразі

$$D^* = \frac{1}{n-1} \sum_{i=1}^n (x_i - m^*)^2$$

незалежні (бо тут m^* залежить від усіх $x_1, x_2, \dots, x_n$).

Але у випадку, коли випадкова величина розподілена нормально або рівномірно, оцінки для дисперсій можна знайти. Можна підрахувати, що

$$D(D^*) = \frac{1}{n} \mu_4 - \frac{n-3}{n(n-1)} D^{*2}.$$

У випадку нормального розподілу:

$$\mu_4 = 3D^2 \quad \text{і тоді} \quad D(D^*) = \frac{2}{n-1} D^{*2},$$

у випадку рівномірного розподілу:

$$\mu_4 = \frac{9}{5} D^2 \quad \text{і тоді} \quad D(D^*) = \frac{4n+6}{5n(n-1)} D^{*2}.$$

Таким чином, у випадку нормального розподілу імовірність

$$P\left\{-\frac{\varepsilon\sqrt{n-1}}{D^*\sqrt{2}} < \frac{D^*-D}{D^*\sqrt{2}}\sqrt{n-1} < \frac{\varepsilon\sqrt{n-1}}{D^*\sqrt{2}}\right\} = \beta.$$

Знову використовуємо функцію Лапласа $\Phi\left(\frac{\varepsilon\sqrt{n-1}}{D^*\sqrt{2}}\right) = \frac{\beta}{2}$ і звідси знаходиться ε .

Існують **точні методи побудови довірчого інтервалу**. При цьому використовуються деякі спеціальні розподіли випадкових величин. Розглянемо чотири таких розподіли.

1. Гамма-розподіл. Функція розподілу випадкової величини має вигляд:

$$F_{\xi}(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{\Gamma(\lambda)} \int_0^x u^{\lambda-1} e^{-u} du, & x \geq 0, \end{cases}$$

тут $\lambda > 0$ – деякий параметр, а $\Gamma(\lambda) = \int_0^{\infty} u^{\lambda-1} e^{-u} du$ – **гамма-функція**.

2. Крім цього, використовується так званий β -розподіл: кажуть, що випадкова величина розподілена за **бета-розподілом**, якщо функція розподілу має вигляд:

$$F_{\xi}(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{B(\lambda_1, \lambda_2)} \int_0^x u^{\lambda_1-1} (1-u)^{\lambda_2-1} du, & 0 \leq x < 1, \\ 1, & x \geq 1, \end{cases}$$

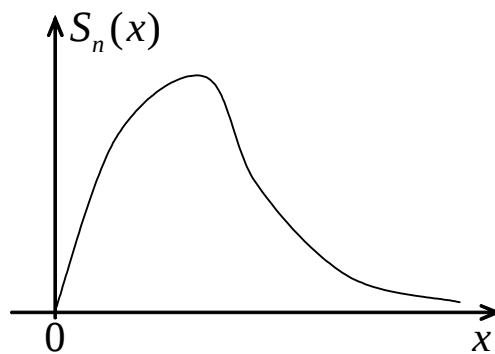
де $\lambda_1 > 0, \lambda_2 > 0, B(\lambda_1, \lambda_2) = \int_0^1 u^{\lambda_1-1} (1-u)^{\lambda_2-1} du$, – **бета-функція**.

3. Якщо $\xi_1, \xi_2, \dots, \xi_n$ – незалежні випадкові величини, розподілені нормально з параметрами $N(0, 1)$, то випадкова величина χ^2 вигляду

$$\chi^2 = \sum_{i=1}^n \xi_i^2$$

розподілена за так званим χ^2 -розподілом **Пірсона*** з n степенями вільності.

Щільність розподілу цієї випадкової величини будемо позначати $S_n(x)$. Зауважимо, що графік $S_n(x)$ має вигляд



4. Якщо ξ і η – незалежні випадкові величини, ξ розподілена нормально з параметрами $N(0, 1)$, а η розподілена за χ^2 -розподілом з n степенями вільності, то випадкова величина $t = \frac{\xi\sqrt{n}}{\sqrt{\eta}}$ розподілена за так званим **розподілом Стьюдента**** з n степенями вільності. Щільність розподілу випадкової величини t записується так:

* Карл Пірсон (Karl (Carl) Pearson, 1857-1936) – англійський математик, статистик, біолог та філософ.

** Вільям Сілі Госсет (псевдонім – Стьюдент) (William Sealy Gosset – Student, 1876-1937) – англійський математик-статистик.

$$f_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

Зауважимо, що при $n = 1$ розподіл Стюдента збігається з розподілом Коші.

Для оцінювання деякого параметра θ використовується певна статистика θ^* . Розподіл випадкової величини θ^* , як правило, залежить від самого параметра θ . Якби розподіл θ^* був відомий, то із співвідношення $P\{|\theta^* - \theta| < \varepsilon\} = \beta$ можна було б знайти ε і побудувати довірчий інтервал. (Вище вже будували такі інтервали наближено, використовуючи нормальність розподілу).

Якщо б вдалося знайти таку функцію від параметра θ^* , а також від точного значення θ і результатів спостережень, розподіл якої не залежить від результатів спостережень (навіть при деяких додаткових умовах), то можна було б точно побудувати довірчий інтервал для параметра, який оцінюємо. Можна довести, що випадкова величина

$$T = \frac{(m^* - m)\sqrt{n}}{\sqrt{D^*}}$$

у випадку нормальності розподілу випадкової величини ξ , за якою спостерігаємо, розподілена за законом Стюдента з $n-1$ степенем вільності. Тут вважаємо, що використано незміщені оцінки

$$m^* = \frac{1}{n} \sum_{i=1}^n x_i, \quad D^* = \frac{1}{n-1} \sum_{i=1}^n (x_i - m^*)^2.$$

Функція розподілу випадкової величини T для різних n затабульована.

Розподіл випадкової величини T не залежить ні від m^* , ні від D^* , а залежить тільки від об'єму вибірки n .

За тих самих умов (ξ розподілена нормально), випадкова величина

$$V = \frac{(n-1)D^*}{D}$$

розподілена за χ^2 -розподілом також з $n-1$ степенем вільності. Використовуючи наведені факти, можна точно побудувати довірчий інтервал.

При побудові довірчого інтервалу для математичного сподівання m використаємо рівність $P\{|m^* - m| < \varepsilon\} = \beta$. Вона еквівалентна такій рівності:

$$P\left\{-\varepsilon \frac{\sqrt{n}}{\sqrt{D^*}} < \frac{(m^* - m)\sqrt{n}}{\sqrt{D^*}} < \varepsilon \frac{\sqrt{n}}{\sqrt{D^*}}\right\} = \beta.$$

Тобто

$$P\left\{|T| < \varepsilon \sqrt{\frac{n}{D^*}}\right\} = \beta.$$

Позначимо $t_\beta = \varepsilon \sqrt{\frac{n}{D^*}}$. Тоді маємо $P\{|T| < t_\beta\} = \beta$. Але таку ймовірність можна виписати, використовуючи щільність розподілу:

$$\int_{-t_\beta}^{t_\beta} f_{n-1}(x) dx = \beta.$$

За таблицями можна знайти t_β . Оскільки функція $f_{n-1}(x)$ парна, то остання рівність така:

$$2 \int_0^{t_\beta} f_{n-1}(x) dx = \beta.$$

Але функція

$$\psi(x) = \int_0^x f_{n-1}(t) dt$$

затабульована, тому за таблицями можна знайти таке x , що $\psi(x) = \frac{\beta}{2}$, враховуючи ще й кількість степенів вільності (які фігурують в таблицях). А після того, як за таблицями буде знайдено t_β , відразу можна обчислити $\varepsilon = t_\beta \sqrt{\frac{D^*}{n}}$. Тоді шуканий інтервал, який вже є точним інтервалом, такий:

$$\left(m^* - t_\beta \sqrt{\frac{D^*}{n}}, m^* + t_\beta \sqrt{\frac{D^*}{n}} \right).$$

Відмінність від аналогічного попереднього запису в тому, що тут t_β знаходимо за таблицями функції Стюдента, а раніше t_β знаходилося наближено за таблицями функції Лапласа.

Точно **побудуємо довірчий інтервал для дисперсії**. Для цього використаємо випадкову величину V . Хочемо побудувати інтервал, в якому з імовірністю β знаходиться D . Тобто хочемо знайти такі числа D_1 і D_2 , щоб $P\{D_1 < D < D_2\} = \beta$.

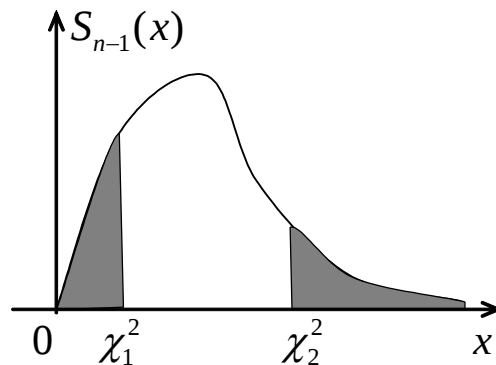
Формально маємо

$$D = \frac{(n-1)D^*}{V}.$$

Скористаємось тут випадковою величиною V . Треба її оцінити (щоб V попало в інтервал (V_1, V_2) і тоді D попаде в якийсь інтервал (D_1, D_2)). Позначимо $\alpha = 1 - \beta$, де β – задане число. Знайдемо числа χ_1^2 і χ_2^2 , такі, що

$$P\{V > \chi_1^2\} = 1 - \frac{\alpha}{2}, \quad P\{V > \chi_2^2\} = \frac{\alpha}{2}.$$

Геометрично це означає, що площі заштрихованих частин підграфіка щільності розподілу $S_{n-1}(x)$ випадкової величини V однакові,



бо вибрано $\frac{\alpha}{2}$. Тут можна було вибрати $1 - \frac{\alpha}{3}$ і $\frac{2}{3}\alpha$. Тоді заштриховані площі були б різні і відрізок (χ_1^2, χ_2^2) перемістився б.

За відомим β (і відповідно α) з виписаних співвідношень за таблицею розподілу Стюдента знаходяться числа χ_1^2 і χ_2^2 . Зрозуміло тепер, що з імовірністю β виконуються нерівності

$$\frac{1}{\chi_2^2} < \frac{1}{V} < \frac{1}{\chi_1^2},$$

а це означає, що з тією ж імовірністю β мають місце нерівності

$$\frac{D^*(n-1)}{\chi_2^2} < \frac{D^*(n-1)}{V} < \frac{D^*(n-1)}{\chi_1^2},$$

тобто

$$P\left\{\frac{D^*(n-1)}{\chi_2^2} < \frac{D^*(n-1)}{V} < \frac{D^*(n-1)}{\chi_1^2}\right\} = \beta.$$

Таким чином, знайшли відповідні числа

$$D_1 = \frac{D^*(n-1)}{\chi_2^2}, \quad D_2 = \frac{D^*(n-1)}{\chi_1^2},$$

а отже, й інтервал (D_1, D_2) .

На практиці побудову довірчого інтервалу для дисперсії здійснюють у такій послідовності:

- 1) задаємо довірчу ймовірність β ;
- 2) знаходимо $\alpha, \frac{\alpha}{2}$;
- 3) використовуючи таблиці Стюдента, одержимо числа χ_1^2 і χ_2^2 ;
- 4) підставляючи у виписані формули одержуємо D_1, D_2 та інтервал (D_1, D_2) .

Такий підхід застосовується при $n \leq 30$. У випадку $n > 30$ можна застосувати інший підхід для інтервальної оцінки, наприклад, середнього квадратичного відхилення σ . Для цього використовуються довірчі інтервали

$$\sigma^*(1-q) < \sigma < \sigma^*(1+q) \quad \text{при } q < 1,$$

$$0 < \sigma < \sigma^*(1+q) \quad \text{при } q > 1,$$

де $q = q(\gamma, n)$ знаходиться за таблицями Стюдента при даних γ і n (таблиця 5 у Додатку В).

Побудуємо довірчий інтервал для ймовірності деякої події A .

Нехай спостерігається деяка подія A , її ймовірність p – невідома, $q = 1 - p$ також невідома. Проводиться спостереження за цією подією. Їй зіставляється випадкова величина ξ :

$$\xi = \begin{cases} 1, & \text{якщо подія відбулася,} \\ 0, & \text{якщо подія не відбулася.} \end{cases}$$

Тобто закон розподілу ξ такий:

ξ	0	1
P	q	p

Фактично спостерігаємо за випадковою величиною ξ . Нехай одержа-

ли такі результати: $x_1, x_2, \dots, x_n$. Тоді частота появи події A при n спостереженнях дорівнює $p^* = \frac{1}{n} \sum_{i=1}^n x_i$. Обчислимо математичне сподівання частоти p^* :

$$\mathbb{E}p^* = \frac{1}{n} \sum_{i=1}^n p = p,$$

тобто частота є незміщеною оцінкою ймовірності p . Крім цього, за теоремою Бернуллі, частота є суттєвою (конзистентною) оцінкою ймовірності p .

Підрахуємо дисперсію Dp^* оцінки p^* :

$$Dp^* = \frac{1}{n^2} \sum_{i=1}^n Dx_i = \frac{1}{n^2} pq \cdot n = \frac{pq}{n}.$$

При $n \rightarrow \infty$ $Dp^* \xrightarrow{n \rightarrow \infty} 0$. Можна показати, що це точна нижня оцінка. А це означає, що частота є ефективною оцінкою ймовірності.

Як же побудувати оцінку для ймовірності (інтервал для ймовірності)? Точно це зробити неможливо. Треба визначити ε , таке, щоб $P\{|p^* - p| < \varepsilon\} = \beta$. Поділимо нерівність $|p^* - p| < \varepsilon$ на $\sqrt{Dp^*}$ і тоді для знаходження ε одержуємо таку рівність:

$$P\left\{\frac{|p^* - p|}{\sqrt{\frac{pq}{n}}} < \varepsilon \sqrt{\frac{n}{pq}}\right\} = \beta.$$

Стосовно випадкової величини $\eta = \frac{|p^* - p|}{\sqrt{\frac{pq}{n}}}$ можна робити при-

пущення, що вона розподілена нормально $N(0, 1)$. Залишається за допомогою функції Лапласа знайти таке $t_\beta = \varepsilon \sqrt{\frac{n}{pq}}$, що $P\{| \eta | < t_\beta\} = \beta$.

Тоді, використовуючи це t_β , запишемо (обчислюємо): $\varepsilon = t_\beta \sqrt{\frac{pq}{n}}$.

Таким чином, виявиться, що з імовірністю β

$$|p^* - p| < \frac{t_\beta}{\sqrt{n}} \sqrt{p(1-p)},$$

тобто

$$P \left\{ |p^* - p| < \frac{t_\beta}{\sqrt{n}} \sqrt{p(1-p)} \right\} = \beta.$$

Фактично шуканий інтервал знайдено:

$$(p^* - \varepsilon, p^* + \varepsilon).$$

Але для знаходження довірчого інтервалу для ймовірності можна не користуватися таблицями, а діяти так: нерівність

$$|p^* - p| < \frac{t_\beta}{\sqrt{n}} \sqrt{p(1-p)}$$

буде виконуватись, якщо

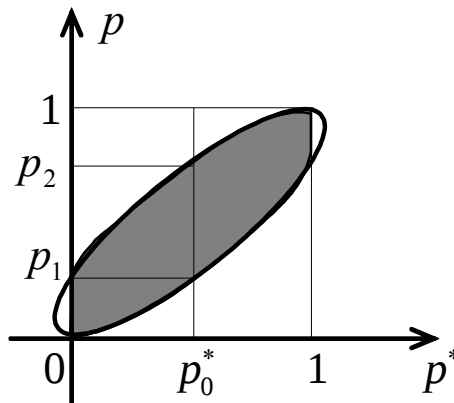
$$(p^* - p)^2 < \frac{t_\beta^2}{n} p(1-p).$$

Розглянемо рівність

$$(p^* - p)^2 = \frac{t_\beta^2}{n} p(1-p).$$

Якщо дослідити криву, яка на площині p^*, p описується цим рівнянням при фіксованому n , то одержується еліпс, який проходить через точки $(0, 0)$ і $(1, 1)$. Допустимі лише точки заштрихованої частина площини, обмеженої еліпсом і квадратом $0 \leq p \leq 1, 0 \leq p^* \leq 1$. Для довільного p_0^* , розв'язуючи квадратне рівняння, одержуються

значення p_1 і p_2 та відповідний інтервал (p_1, p_2) .



У випадку векторної випадкової величини (наприклад, двовимірної (ξ, η)) результати спостережень мають вигляд $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Як оцінки параметрів використовуються векторні оцінки

$$\bar{m}^* = (m_x^*, m_y^*), \quad \text{де} \quad m_x^* = \frac{1}{n} \sum_{i=1}^n x_i, \quad m_y^* = \frac{1}{n} \sum_{i=1}^n y_i.$$

В результаті будується довірчий прямокутник для оцінки математичного сподівання.

Для оцінки дисперсії використовуються такі співвідношення:

$$D_x^* = \left[\frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)^2 \right] \cdot \frac{n}{n-1}, \quad D_y^* = \left[\frac{1}{n} \sum_{i=1}^n (y_i - m_y^*)^2 \right] \cdot \frac{n}{n-1}.$$

Можна записати оцінки і для інших характеристик векторної випадкової величини. Так, для змішаних центральних моментів треба використати оцінку

$$K_{xy}^* = \left[\frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)(y_i - m_y^*) \right] \cdot \frac{n}{n-1},$$

для коваріаційної матриці оцінкою є матриця

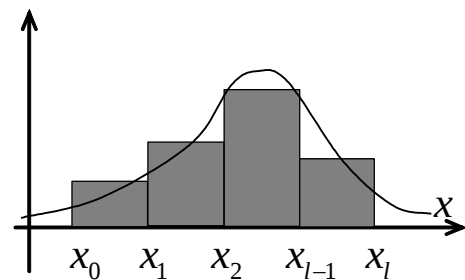
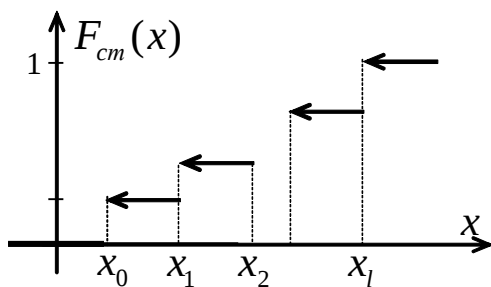
$$\begin{pmatrix} D_x^* & K_{xy}^* \\ K_{xy}^* & D_y^* \end{pmatrix}.$$

§5. Критерії згоди

Припустимо, що при попередній обробці результатів спостережень побудовано варіаційний ряд

$[x_0, x_1)$	$[x_1, x_2)$	$\dots$	$[x_{l-1}, x_l]$
p_1^*	p_2^*	$\dots$	p_l^*

і визначені деякі точкові оцінки параметрів розподілу m^* , D^* . За варіаційним рядом легко побудувати статистичну функцію розподілу $F_{cm}(x)$ або гістограму. Нехай вони мають такий вигляд:



За отриманими результатами можна зробити попередній висновок про тип розподілу випадкової величини ξ . В останньому прикладі можна висловлювати припущення, що функція розподілена нормально. Попередній висновок про розподіл роблять після побудови так званої згладжуючої кривої, яка отримується за результатами обробки варіаційного ряду.

Припустимо, що висловлена деяка гіпотеза H про розподіл випадкової величини. Наприклад, H – „функція розподілу $F_\xi(x)$ випадкової величини має вигляд...” або „щільність розподілу $f_\xi(x)$ має вигляд...” або „випадкова величина ξ – розподілена за таким законом:”

ξ	x_1	$\dots$	x_n
P	p_1	$\dots$	p_n

Постає питання про те, наскільки ця гіпотеза узгоджується з результатами експериментів (спостережень)? Отже, маємо частоти $p_1^*, p_2^*, \dots, p_l^*$ результатів експериментів. Використовуючи гіпотетичну функцію розподілу можна підрахувати ймовірності $p_1, p_2, \dots, p_l$ того, що випадкова величина, яка розподілена за гіпотетичним законом, набуває значення з відповідного проміжку: $p_i = P_F \{x_{i-1} \leq \xi < x_i\}$. Треба порівняти між собою два набори $p_1^*, p_2^*, \dots, p_l^*$ і $p_1, p_2, \dots, p_l$. Якщо вони, в якомусь розумінні, мало відрізняються, то можна прийняти гіпотезу. Можна було б розглядати $\sum_{i=1}^l (p_i^* - p_i)^2$. Але більш природно ввести ще якісь вагові коефіцієнти C_i і розглядати $U = \sum_{i=1}^l C_i (p_i^* - p_i)^2$. Пірсон запропонував вибирати коефіцієнти C_i у вигляді $C_i = \frac{n}{p_i}$ і, відповідно, міру розходження (відмінності) результатів статистичної обробки і результатів за гіпотетичним розподілом розглядати у вигляді

$$U = n \sum_{i=1}^l \frac{(p_i^* - p_i)^2}{p_i} = \chi^2.$$

Зрозуміло, що міра відмінності (U) є випадковою величиною. Її розподіл залежить від розподілу досліджуваної випадкової величини (бо частоти p_i^* є випадковими величинами). Пірсон довів, що розподіл випадкової величини χ^2 слабо залежить від розподілу досліджуваної випадкової величини, тобто при $n \rightarrow \infty$ розподіл випадкової величини χ^2 (взагалі кажучи, залежний від n : χ_n^2) прямує до розподілу незалежного від n . Граничний розподіл залежить тільки від кількості інтервалів (від l).

Пірсон довів, що граничним розподілом випадкової величини χ^2 є так званий **χ^2 -розподіл (розподіл Пірсона)** із кількістю степенів вільності, тісно пов'язаною із кількістю інтервалів l (від n залежності нема). А саме: розподіл числа χ^2 є χ^2 -розподілом Пірсона з кількістю степенів вільності $r = l - s$, де l – кількість інтервалів, s – кількість зв'язків, накладених на гіпотетичну функцію розподілу. Що це за зв'язки? Один зв'язок завжди наявний і має вигляд: $p_1^* + p_2^* + \dots + p_l^* = 1$ – сума частот дорівнює одиниці, тобто $s \geq 1$.

На гіпотетичну функцію розподілу випадкової величини ξ може накладатися, наприклад, такий зв'язок: $E\xi = m^*$. Тоді вже маємо $s \geq 2$. Крім цього, може накладатися ще умова $D\xi = D^*$, і тоді вже $s \geq 3$. Наприклад, якщо видно, що закон розподілу нормальний, то щільність розподілу

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

і виникає запитання, які значення σ і m ? Можна прийняти $\sigma = \sigma^*$, $m = m^*$, тобто

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

і тоді фактично є три зв'язки $s = 3$, тому кількість степенів вільності $r = l - 3$.

Пірсоном складено таблиці відповідної функції розподілу. Для практичного користування складено таблиці ймовірностей такого типу: $P\{U \geq \chi^2\}$ (таблиця 6 у Додатку В). За таблицями можна визначити ймовірність того, що випадкова величина U , яка розподілена за законом Пірсона, набуває значення не менше, ніж обчислена міра

розбіжності χ^2 . Якщо ця ймовірність мала, то гіпотезу про розподіл випадкової величини ξ слід відкинути як таку, що суперечить результатам спостережень (експериментів).

Пояснення: після того, як підраховували χ^2 (одержали конкретне значення), за таблицями підраховуємо ймовірність того, що розбіжність U може бути більшою. Якщо ця ймовірність велика, то гіпотезу можна прийняти і навпаки.

На практиці вважають: якщо ця ймовірність менша за 0,1, то гіпотезу треба відкидати, бо вона суперечить результатам експериментів.

Якщо ця ймовірність дуже велика, то треба експеримент повторити і переконатися, чи не помилкові результати.

Якщо ця ймовірність достатньо велика (при $p = 0,2$ і більше), то це не означає, що гіпотеза правильна, а лише те, що гіпотеза не суперечить результатам експерименту і цією гіпотезою можна користуватись.

Алгоритм перевірки статистичної гіпотези:

1. Висунути гіпотезу, тобто це може бути гіпотеза про функцію $F_\xi(x)$ або щільність $f_\xi(x)$ розподілу.
2. Підрахувати за $F_\xi(x)$ чи $f_\xi(x)$ ймовірності $p_1, p_2, \dots, p_l$.
3. Підрахувати число χ^2 .
4. За таблицями знайти $P\{U \geq \chi^2\}$.
5. Зробити висновок: якщо $P\{U \geq \chi^2\} \leq 0,1$, то гіпотеза відкидається.

Часто для обчислення U використовують таку формулу (вносимо у формулі для U n під знак суми):

$$U = \sum_{i=1}^l \frac{(n_i - np_i)^2}{np_i},$$

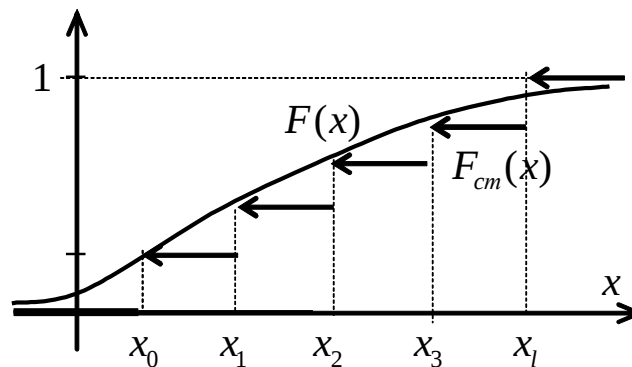
де n_i – кількість результатів експериментів, які потрапили в i -ий інтервал. Потрібно, щоб $n_i \geq 5$, якщо $n_i < 5$, то інтервали треба об'єднати.

Зауваження: χ^2 -розподіл справедливий лише для нульової гіпотези. Для альтернативної гіпотези розподіл статистики може відрізнятися від χ^2 -розподілу. Тому при перевірці нульової гіпотези χ^2 -розподіл корисний і застосовний, тоді як для альтернативної гіпотези цей розподіл може бути неправильним.

Існує багато аналогічних критеріїв. Широко розповсюджений **критерій згоди Колмогорова**. У ньому використовується гіпотетична та статистична функції розподілу. Обчислюється число

$$D = \max |F(x) - F_{cm}(x)|$$

(тут максимум на $\mathbb{R}$). Це число D можна знайти або оцінити за статистичною функцією розподілу.



Колмогоров довів, що випадкова величина U , яка дорівнює $U_n = D\sqrt{n}$ при $n \rightarrow \infty$, слабо залежить від розподілу досліджуваної випадкової величини. Якщо б не була функція розподілу $F(x)$ неперервної випадкової величини, ймовірність $P\{U_n \geq \lambda\}$ нерівності $D\sqrt{n} \geq \lambda$ прямує, при $n \rightarrow \infty$, до границі

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda^2}.$$

Ним складено таблицю ймовірностей $P(\lambda)$, яка наведена у Додатку В (таблиця 7).

Якщо ймовірність $P(\lambda)$ мала (менша рівня значущості), то гіпотезу про розподіл випадкової величини за запропонованим законом варто відхилити. Якщо ж ймовірність $P(\lambda)$ досить велика (більша рівня значущості), то гіпотеза про запропонований закон розподілу випадкової величини визнається такою, що не суперечить результатам експерименту. Розбіжності між $F_{cm}(x)$ і $F(x)$ вважаються несуттєвими і можуть бути пояснені випадковими факторами.

Алгоритм застосування критерію Колмогорова:

1. Підрахувати D .
2. Обчислити $\lambda = D\sqrt{n}$.
3. Знайти $P(\lambda)$ за таблицями значень функції $P(\lambda) = P\{U_n \geq \lambda\}$.
4. Гіпотеза відхиляється, якщо $P(\lambda)$ менше рівня значущості, а якщо $P(\lambda)$ більше рівня значущості, то гіпотезу варто визнати такою, що не суперечить експерименту.

Критерій Колмогорова простіший за критерій χ^2 Пірсона, тому його охоче застосовують на практиці. Проте цей критерій можна застосовувати лише в тих випадках, коли відомий гіпотетичний розподіл $F(x)$ разом з усіма його параметрами. Такі ситуації порівняно рідко трапляються на практиці.

Критерій Колмогорова абсолютно не залежить від накладених зв'язків (можливо, тільки опосередковано), а це не дуже добре. Це так званий критерій другого рангу, є ще інші критерії першого рангу, нульового рангу.

§6. Принцип найбільшої правдоподібності. Метод найменших квадратів згладжування результатів експерименту

Нехай проводяться вимірювання деякого фактора Y , який залежить від аргументу x (x не є випадковою величиною). Тобто при конкретних значеннях аргументу x вимірюється значення Y . Якась залежність Y від x існує $y = \varphi(x)$, але вона (залежність) невідома. Нехай вимірювання при $x_1, x_2, \dots, x_n$ дали такі результати:

x_1	x_2	...	x_n
y_1	y_2	...	y_n

У зв'язку з тим, що при вимірюваннях, як правило, робляться похибки, числа $y_1, y_2, \dots, y_n$ слід вважати реалізаціями випадкових величин $Y_1, Y_2, \dots, Y_n$ (при інших замірах для x_1 можна одержати інші значення y_1 і т.д.).

Згідно з граничними теоремами, при однакових умовах вимірювання (спостереження) слід вважати, що випадкові величини $Y_1, Y_2, \dots, Y_n$ розподілені нормально. Їх математичні сподівання повинні збігатися з точним значенням, тобто $EY_1 = \varphi(x_1)$, $EY_2 = \varphi(x_2), \dots$, $EY_n = \varphi(x_n)$. Будемо вважати, що вимірювання при всіх x однаково точні (проводяться одним приладом і т.ін.), а це означає, що відхилення від точного значення однакові, тобто $DY_1 = DY_2 = \dots = DY_n = \sigma^2$. Тоді випадкові величини Y_i , $i = \overline{1, n}$, розподілені за таким законом: щільність розподілу записується у вигляді

$$f_i(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_i - \varphi(x_i))^2}{2\sigma^2}}.$$

Вважаємо, що всі вимірювання незалежні. Тоді ймовірність попадання y_i , для всіх $i = \overline{1, n}$, на відрізок довжиною Δ така:

$$P\left\{y_i \in \left[y_i - \frac{\Delta}{2}, y_i + \frac{\Delta}{2}\right]\right\} \approx f_i(y_i) \cdot \Delta,$$

і, оскільки вимірювання незалежні, то одержуємо добуток імовірностей

$$P\left\{y_i \in \left[y_i - \frac{\Delta_i}{2}, y_i + \frac{\Delta_i}{2}\right], i = \overline{1, n}\right\} \cong \prod_{i=1}^n f_i(y_i) \cdot K = e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \varphi(x_i))^2} \cdot K_1,$$

де $K_1 = K \cdot \frac{1}{(\sigma\sqrt{2\pi})^n}$, K – стала. Як підібрати функцію φ , щоб ця

ймовірність була найбільшою (щоб була найбільш правдоподібною подія – попадання на відповідні відрізки всіх y_i)? Треба, щоб показник степеня був найменшим, тобто

$$\sum_{i=1}^n (y_i - \varphi(x_i))^2 \rightarrow \min_{\Phi}.$$

Тут мінімум по всіх класах функцій φ , фактично по всіх коефіцієнтах.

Найчастіше загальний вигляд функції φ відомий. Як правило, вибирають

$$\varphi_1(x) = ax + b$$

або

$$\varphi_2(x) = ax^2 + bx + c$$

або

$$\varphi_3(x) = A \cdot e^{Bx}$$

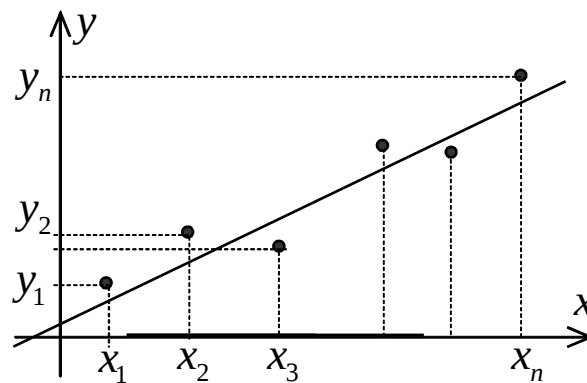
і т.д.

Невідомі лише відповідні коефіцієнти (параметри). Їх і знаходять. Розглянемо лише функцію $\varphi_1(x)$. Фактично маємо задачу мінімізації

$$\sum_{i=1}^n [y_i - \varphi(x_i, a, b, c, \dots)]^2 \rightarrow \min_{a, b, c, \dots}.$$

Для випадку $\varphi_1(x)$ вона набуває вигляду

$$\sum_{i=1}^n [y_i - ax_i - b]^2 \rightarrow \min_{a, b} \quad a \in \mathbb{R}, b \in \mathbb{R}.$$



Прирівнюємо похідні по a та b до нуля:

$$\begin{cases} \sum_{i=1}^n [y_i - ax_i - b] \cdot x_i = 0, \\ \sum_{i=1}^n [y_i - ax_i - b] = 0, \end{cases} \quad \begin{cases} \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - bn = 0, \end{cases}$$

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i + bn = \sum_{i=1}^n y_i, \end{cases}$$

Поділимо на n обидві рівності останньої системи:

$$\begin{cases} a \frac{1}{n} \sum_{i=1}^n x_i^2 + b \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n x_i y_i, \\ a \frac{1}{n} \sum_{i=1}^n x_i + b = \frac{1}{n} \sum_{i=1}^n y_i. \end{cases}$$

Позначимо: $\alpha_{10}^* = m_x^* = \frac{1}{n} \sum_{i=1}^n x_i$, де x_i – не випадкові і тому α_{10}^* це

є просто середнє арифметичне; можна трактувати α_{10}^* як аналог змішаного початкового моменту, тут 1 – показник степеня x_i ; 0 – показник y_i ;

$$\alpha_{20}^* = \frac{1}{n} \sum_{i=1}^n x_i^2, \quad \alpha_{11}^* = \frac{1}{n} \sum_{i=1}^n x_i y_i, \quad \alpha_{01}^* = \frac{1}{n} \sum_{i=1}^n y_i.$$

Тоді система рівнянь набуває такого вигляду:

$$\begin{cases} a\alpha_{20}^* + b\alpha_{10}^* = \alpha_{11}^*, \\ a\alpha_{10}^* + b = \alpha_{01}^*. \end{cases}$$

Звідси $b = \alpha_{01}^* - a\alpha_{10}^*$. Підставляємо це значення у перше рівняння. Одержимо

$$a = \frac{\alpha_{11}^* - \alpha_{10}^* \cdot \alpha_{01}^*}{\alpha_{20}^* - (\alpha_{10}^*)^2}.$$

Таким чином, отримуємо:

$$y = ax + b,$$

тут a за формою запису є відношенням **незміщених** оцінок, а саме

$a = \frac{K_{11}^*}{D_x^*}$, де K_{11}^* – аналог коваріаційного моменту x, y , а D_x^* – аналог

для оцінки дисперсії; $b = m_y^* - am_x^*$.

Отже, маємо:

$$y = m_y^* + \frac{K_{11}^*}{D_x^*}(x - m_x^*)$$

або

$$y - m_y^* = \frac{K_{11}^*}{D_x^*}(x - m_x^*).$$

Аналогічно, у випадку функції

$$y = \varphi_2(x) = ax^2 + bx + c,$$

коефіцієнти a, b, c шукають з умови

$$\sum_{i=1}^n [y_i - ax_i^2 - bx_i - c]^2 \rightarrow \min_{a,b,c}.$$

Для знаходження a, b, c отримується система лінійних рівнянь

$$\begin{cases} a \cdot \alpha_{40}^* + b \cdot \alpha_{30}^* + c \cdot \alpha_{20}^* = \alpha_{21}^*, \\ a \cdot \alpha_{30}^* + b \cdot \alpha_{20}^* + c \cdot \alpha_{10}^* = \alpha_{11}^*, \\ a \cdot \alpha_{20}^* + b \cdot \alpha_{10}^* + c = \alpha_{01}^*. \end{cases}$$

Структура цієї системи така сама, як у випадку $\varphi_1(x)$.

Якщо функція $\varphi(x, a, b, c, \dots)$ нелінійна, то відповідна система рівнянь виявляється досить складною.

§7. Згладжування результатів спостережень за двовимірною випадковою величиною.

Регресія. Лінійна регресія

У попередньому параграфі x_i вважалися детермінованими (невипадковими). Тут припустимо, що спостерігається векторна випадкова величина (ξ, η) .

Реалізації випадкових величин компонент вектора такі:

$$\xi - x_1, x_2, \dots, x_n, \quad \eta - y_1, y_2, \dots, y_n,$$

тобто реалізаціями, у результаті спостережень, будуть пари

$$(x_i, y_i), \quad i = \overline{1, n}.$$

Як і для одновимірної випадкової величини, треба провести попередню обробку результатів спостережень. Найчастіше використовується аналог варіаційного ряду. Для кожної з випадкових величин можна будувати варіаційний ряд. Не обов'язково з однаковим кроком, не обов'язково з однаковою кількістю інтервалів. У результаті одержується двовимірна таблиця

$\xi \backslash \eta$	$[x_0, x_1)$ 1	$[x_1, x_2)$ 2	$[x_2, x_3)$ 3	$[x_3, x_4]$ 4
$[y_0, y_1)$ 1	n_{11}			
$[y_1, y_2)$ 2		n_{22}	n_{23}	
$[y_2, y_3)$ 3			n_{33}	n_{34}
$[y_3, y_4)$ 4				n_{44}
$[y_4, y_5]$ 5				n_{45}

де, наприклад, n_{23} – кількість результатів експерименту, що попали у клітинку, тобто $[x_2, x_3) \times [y_1, y_2)$. Якщо клітинка незаповнена, то

вважається, що відповідна $n_{ij} = 0$. За таблицею можна робити попередні висновки про зв'язок між випадковими величинами. Крім того, можна підрахувати оцінки $\mathbb{E}\xi \approx m_\xi^*$, $\mathbb{E}\eta \approx m_\eta^*$, $D\xi \approx D_\xi^*$, $D\eta \approx D_\eta^*$. Для двовимірної випадкової величини важливі характеристики аналогії змішаних моментів $K_{\xi\eta} \approx K_{\xi\eta}^*$, $\rho(\xi, \eta) \approx \rho_{\xi\eta}^* = \frac{K_{\xi\eta}^*}{\sigma_\xi^* \cdot \sigma_\eta^*}$.

Для векторних випадкових величин важливі характеристики – умовні розподіли однієї випадкової величини при фіксованому значенні іншої випадкової величини. Найпростішою характеристикою такого плану є умовні математичні сподівання. При обробці результатів спостережень над векторними випадковими величинами умовні математичні сподівання відіграють основну роль.

Нехай задана система двох дискретних випадкових величин (ξ, η) і ймовірності $p_{ij} = P(\xi = x_i \cap \eta = y_j)$. В загальному випадку, якщо ξ та η залежні, то, за теоремою добутку ймовірностей, можна записати

$$p_{ij} = P(\xi = x_i) \cdot P(\eta = y_j \mid \xi = x_i) = p_{i*} \cdot p_{x_i}(y_j).$$

Аналогічно можна одержати і таку формулу: $p_{ij} = p_{j*} \cdot p_{y_j}(x_i)$.

Вирази $p_{x_i}(y_j)$ і $p_{y_j}(x_i)$ визначають умовні закони розподілу випадкових величин η та ξ , тобто розподіли однієї випадкової величини, знайдені за умови, що інша випадкова величина набула визначене значення. З наведених вище формул випливає, що

$$p_{x_i}(y_j) = \frac{p_{ij}}{p_{i*}} = \frac{p_{ij}}{\sum_{j=1}^n p_{ij}}, \quad p_{y_j}(x_i) = \frac{p_{ij}}{p_{j*}} = \frac{p_{ij}}{\sum_{i=1}^m p_{ij}}.$$

Означення 1. Умовним математичним сподіванням дискретної випадкової величини ξ при фіксованому значенні η називається сума добутків можливих значень ξ на їхні умовні ймовірності

$$\mathbb{E}(\xi \mid \eta = y_j) = \sum_{i=1}^m x_i \cdot p_{y_j}(x_i).$$

Аналогічно вводиться поняття щільності умовного розподілу, для системи неперервних випадкових величин у такому вигляді:

$$f_{\xi|\eta}(x \mid y) = \frac{f_{\xi\eta}(x, y)}{f_{\eta}(y)}.$$

Тоді можна сформулювати означення умовного математичного сподівання.

Означення 2. Умовним математичним сподіванням абсолютно неперервної випадкової величини ξ при фіксованому значенні η називається інтеграл від добутку змінної на щільність умовного розподілу

$$\mathbb{E}(\xi \mid \eta = y) = \int_{-\infty}^{+\infty} x \cdot f_{\xi|\eta}(x \mid y) dx.$$

Функціональний зв'язок між умовним математичним сподіванням

$$Y = \mathbb{E}(\eta \mid \xi = x)$$

випадкової величини η при фіксованому значенні випадкової величини $\xi = x$ і самим x називається **регресією η на ξ** .

Функціональний зв'язок між умовним математичним сподіванням

$$X = \mathbb{E}(\xi \mid \eta = y)$$

і самим y називається **регресією ξ на η** .

Якщо припустити, що регресія лінійна, то форма зв'язку має вигляд

$$Y = Ax + B \quad - \text{регресія } \eta \text{ на } \xi$$

або

$$X = A'y + B' \quad - \text{регресія } \xi \text{ на } \eta.$$

Відшукування (побудова) лінійної регресії зводиться до побудови методом найменших квадратів зв'язків між $m_{\eta|x}$ і x та $m_{\xi|y}$ і y (тут $m_{\eta|x}$, $m_{\xi|y}$ – математичні сподівання, які записані вище). Тепер при такій побудові всі моменти α_{ij}^* , які фігурували у попередньому параграфі мають імовірнісний зміст. Отже, рівняння прямих регресії мають вигляд рівняння прямої регресії η на ξ :

$$Y - m_{\eta}^* = \frac{K_{\xi\eta}^*}{D_{\xi}^*} (X - m_{\xi}^*),$$

рівняння прямої регресії ξ на η :

$$X - m_{\xi}^* = \frac{K_{\xi\eta}^*}{D_{\eta}^*} (Y - m_{\eta}^*).$$

Це рівняння різних прямих. Якщо мати на увазі, що

$$\rho(\xi, \eta) = \frac{K_{\xi\eta}}{\sigma_{\xi} \cdot \sigma_{\eta}}, \quad \text{а тому} \quad \rho_{\xi\eta}^* = \frac{K_{\xi\eta}^*}{\sigma_{\xi}^* \cdot \sigma_{\eta}^*},$$

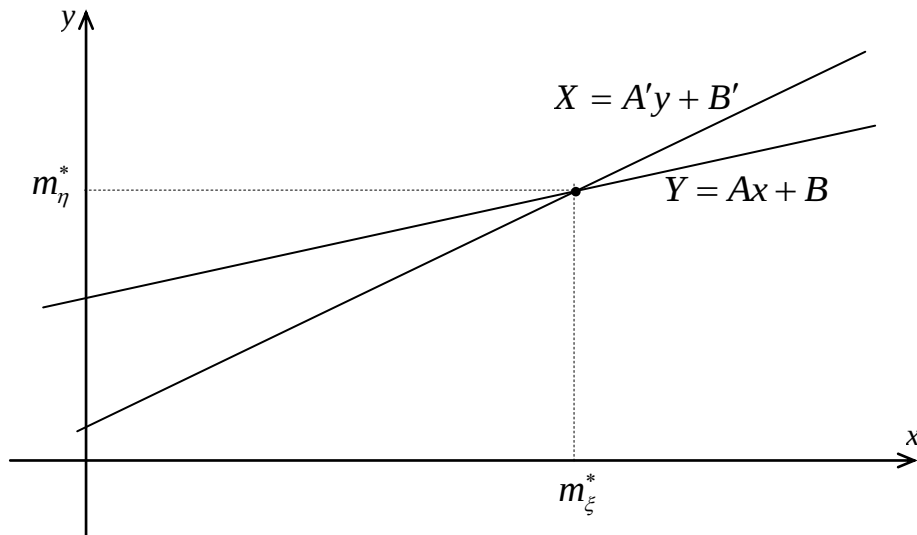
то рівняння прямих регресії записуються у такому вигляді:

$$Y - m_{\eta}^* = \frac{\sigma_{\eta}^*}{\sigma_{\xi}^*} \rho_{\xi\eta}^* (X - m_{\xi}^*), \quad (1)$$

$$X - m_{\xi}^* = \frac{\sigma_{\xi}^*}{\sigma_{\eta}^*} \rho_{\xi\eta}^* (Y - m_{\eta}^*). \quad (2)$$

Природно, що за регресією η на ξ можна здійснювати прогноз значень η при фіксованих значеннях ξ . За рівнянням (2) можна здійснювати прогноз ξ при заданому η .

З рівнянь (1), (2) зрозуміло, що прямі проходять через точку (m_ξ^*, m_η^*) .



Якщо коефіцієнт кореляції більший нуля, то зв'язок називається **прямим** (при збільшенні ξ в середньому збільшується і η , і навпаки, – при збільшенні η збільшується ξ).

Якщо коефіцієнт кореляції менший нуля, то зв'язок називається **оберненим**.

Якщо коефіцієнт кореляції близький до одиниці, то зв'язок – **сильний**.

Якщо коефіцієнт кореляції близький до нуля, то зв'язок – **слабкий**.

Після побудови зв'язку може виникати питання про його достовірність. Тут все вирішується, як і раніше, через довірчі інтервали. Так для зв'язку $Y = AX + B$ довірчий інтервал для коефіцієнта A такий:

$$A^* - t_\beta \frac{\sigma_\eta^*}{S_1 \sqrt{n-2}} \leq A \leq A^* + t_\beta \frac{\sigma_\eta^*}{S_1 \sqrt{n-2}},$$

де

$$A^* = \frac{\sigma_\eta^*}{\sigma_\xi^*} \rho_{\xi\eta}^*, \quad S_1 = \frac{1}{n} \sum_{i=1}^n (x_i - m_\xi^*)^2,$$

ймовірність t_β знаходиться з рівняння

$$P\{|\tau| > t_\beta\} = \beta,$$

де τ – випадкова величина, розподілена за законом Стюдента з $(n-2)$ -а степенями вільності.

Аналогічно виписується довірчий інтервал для коефіцієнта B :

$$B^* - t_\beta \frac{\sigma_\eta^*}{\sqrt{n-2}} \leq B \leq B^* + t_\beta \frac{\sigma_\eta^*}{\sqrt{n-2}},$$

де

$$B^* = A^* m_\xi^* + m_\eta^*.$$

Подальше узагальнення може полягати у побудові нелінійної регресії. Або ж, залишаючись з лінійною регресією (зв'язком), розглядати тривимірні векторні випадкові величини. Потім можна ще й для тривимірних векторів розглядати нелінійні зв'язки.

РОЗДІЛ 5

ЕЛЕМЕНТИ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ (ФУНКЦІЙ)

§1. Основні поняття

Нехай задано ймовірнісний простір $(\Omega, \mathcal{F}, P)$.

Означення 1. *Випадковим процесом (випадковою функцією) на заданому ймовірнісному просторі $(\Omega, \mathcal{F}, P)$ називають сімейство випадкових величин $x_t(\omega)$, вимірних відносно σ -алгебри $\mathcal{F}$ для заданих (деяких, не будь-яких) значень $t, t \in \mathbb{T}, \omega \in \Omega$.*

Інші позначення випадкових процесів: $x_t, x(t), x = \{x_t, t \in \mathbb{T}\}$. Як правило, множина значень параметра t повинна задаватись, її інтерпретують як час. Розрізняють випадкові процеси з дискретним часом і неперервним часом:

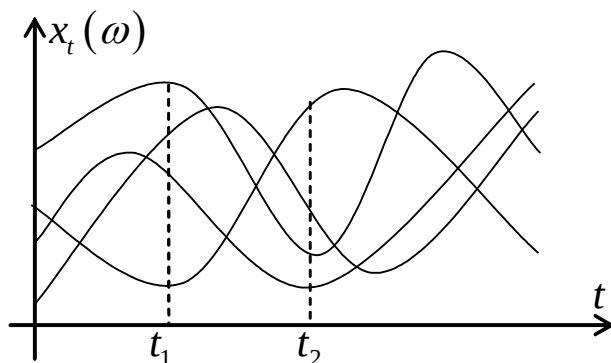
- ✓ якщо множина $\mathbb{T}$ не більше, ніж злічена, то кажуть, що розглядається **процес з дискретним часом**; тоді $\mathbb{T} = \mathbb{N} = \{1, 2, 3, \dots\}$ або $\mathbb{T} = \mathbb{N} \cup \{0\}$;
- ✓ якщо часовий параметр неперервний, то кажуть, що розглядається **процес з неперервним часом** і тоді $\mathbb{T} = [a, b]$ або $\mathbb{T} = [a, +\infty)$ або $\mathbb{T} = \mathbb{R}_+$.

Процеси також розрізняють за множиною значень.

Якщо випадкова величина $x_t(\omega)$ дискретна при кожному t , то кажуть, що розглядається процес з дискретними значеннями. Якщо ж

$x_t(\omega)$ належить класу абсолютно неперервних випадкових величин, то кажуть, що розглядається процес з неперервними значеннями.

Множина значень процесу вважається однаковою для всіх t і позначається X : $x_t(\omega) \in X$, X – фазовий простір. При фіксованому $\omega \in \Omega$ одержується якась функція від t , це може бути інша випадкова величина, розподілена зовсім за іншим законом.



Випадковий процес слід розуміти як сімейство функцій відразу.

Інформація про векторну випадкову величину задається функцією сумісного розподілу компонент.

Вважається, що процес $x_t(\omega)$ заданий, якщо для будь-яких $t_1 < t_2 < t_3 < \dots < t_n$ і будь-якого n ($t_i \in \mathbb{T}$, $i = \overline{1, n}$) задано функцію сумісного розподілу

$$F_{t_1 t_2 \dots t_n}(y_1, y_2, \dots, y_n) = P\{\omega: x_{t_1} < y_1, x_{t_2} < y_2, \dots, x_{t_n} < y_n\}.$$

Причому функції сумісних розподілів узгоджені між собою у тому розумінні, що, наприклад,

$$F_{t_1 t_2 t_3 \dots t_n}(y_1, \infty, y_3, \dots, y_n) = F_{t_1 t_3 \dots t_n}(y_1, y_3, \dots, y_n).$$

Отже, повністю задати випадковий процес при допомозі сімейства функцій розподілу практично дуже складно (неможливо). Тому серед випадкових процесів виділяються класи таких процесів, які задовольняють ще деякі додаткові умови, а ці умови дають можли-

вість будувати сімейство функцій розподілу. Серед таких процесів найпростіші:

- процеси з незалежними значеннями,
- процеси з незалежними приростами,
- однорідні процеси,
- стаціонарні процеси,
- марковські процеси.

Означення 2. Процес x_t називається **процесом з незалежними значеннями**, якщо для всякого набору $t_1 < t_2 < t_3 < \dots < t_n$ випадкові величини $x_{t_1}, x_{t_2}, \dots, x_{t_n}$ – незалежні.

У цьому випадку

$$\begin{aligned} F_{t_1 t_2 \dots t_n}(y_1, y_2, \dots, y_n) &= P\{\omega: x_{t_1} < y_1, x_{t_2} < y_2, \dots, x_{t_n} < y_n\} = \\ &= \prod_{i=1}^n P\{\omega: x_{t_i} < y_i\} = \prod_{i=1}^n F_{t_i}(y_i). \end{aligned}$$

Тому сімейство функцій сумісного розподілу задається як добуток функцій одновимірних розподілів, тобто все зводиться до задання сімейства одновимірних розподілів.

Означення 3. Невипадкова функція $m(t)$ називається **математичним сподіванням випадкового процесу** x_t , якщо для кожного фіксованого $t \in \mathbb{T}$ математичне сподівання $\mathbb{E}x_t = m(t)$.

(тут зліва математичне сподівання випадкової величини x_t , бо t фіксоване).

Означення 4. Невипадкова функція $D(t) = \sigma^2(t)$ називається **дисперсією випадкового процесу** x_t , якщо для кожного фіксованого $t \in \mathbb{T}$ дисперсія $Dx_t = D(t)$.

Можливі й інші записи:

$$Dx_t = D(t) = \mathbb{E}(x_t - m(t))^2 = \mathbb{E}x_t^2 - (m(t))^2.$$

Означення 5. Коваріаційною функцією випадкового процесу x_t називається функція

$$B(t, s) = \text{cov}(x_t, x_s) = \mathbb{E}(x_t - m(t))(x_s - m(s)), \quad t \in \mathbb{T}, \quad s \in \mathbb{T}.$$

Означення 6. Функція $\rho(t, s) = \frac{B(t, s)}{\sigma(t)\sigma(s)}$ називається **кореляційною** або **автокореляційною функцією** випадкового процесу x_t .

Важливим процесом є такий процес, який повністю визначається функціями $m(t)$, $D(t)$ і $B(t, s)$. Більше того, якщо

$$m(t) = 0, \quad D(t) = \sigma^2(t) = \text{const}, \quad B(t, s) = b(t - s) = b(s - t),$$

то процес x_t називається **стаціонарним у широкому розумінні**.

Тут через $b(t)$ позначено кореляційну функцію стаціонарного випадкового процесу:

$$b(t) = \text{cov}(x_{s+t}, x_s).$$

Зауважимо, що $B(t, t) = D(t)$. Тому серед умов стаціонарності є третя умова, а друга впливає одразу, першої можна добитися, перейшовши до центрованих величин.

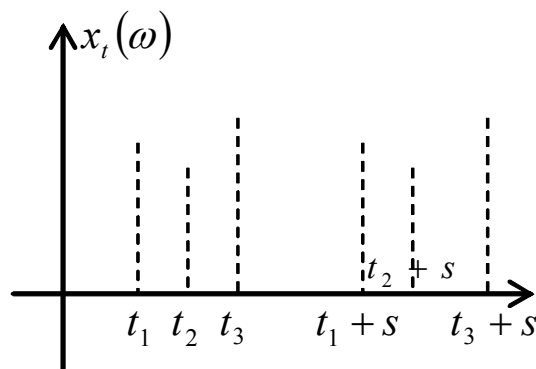
Параметрами (функціями) $m(t)$, $D(t)$, $B(t, s)$ або $b(t)$ описується такий процес, що функція розподілу $F_{t_1 t_2}(y_1, y_2)$ є функцією нормального двовимірного розподілу.

Часто зберігаючи основні властивості будь-якого процесу (моменти до другого порядку) всякий процес замінюють процесом з нормальним двовимірним розподілом. Таке вивчення процесу $x_t(\omega)$ називається вивченням процесу **у широкому розумінні**.

Процес $x_t(\omega)$ називається **стаціонарним у вузькому розумінні**, якщо для будь-якого n і для будь-яких $t_1 < t_2 < t_3 < \dots < t_n$ функція розподілу $F_{t_1+s, t_2+s, \dots, t_n+s}(y_1, y_2, \dots, y_n)$ співпадає з функцією розподілу $F_{t_1, t_2, \dots, t_n}(y_1, y_2, \dots, y_n)$:

$$F_{t_1+s, t_2+s, \dots, t_n+s}(y_1, y_2, \dots, y_n) = F_{t_1, t_2, \dots, t_n}(y_1, y_2, \dots, y_n)$$

(кажуть, що функція сумісного розподілу не залежить від зміщення часу).



Якщо випадковий процес стаціонарний у вузькому розумінні, то він стаціонарний і в широкому розумінні. Навпаки, взагалі кажучи, неправильно.

Означення 7. Процес x_t називається **процесом з незалежними приростами**, якщо для будь-якого набору $t_1 < t_2 < t_3 < \dots < t_n$ і для будь-якого n , випадкові величини

$$x_{t_1}, \quad x_{t_2} - x_{t_1}, \quad x_{t_3} - x_{t_2}, \dots, \quad x_{t_n} - x_{t_{n-1}}$$

незалежні.

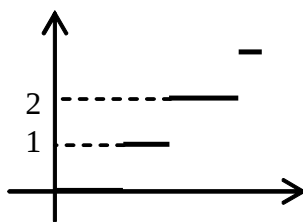
Означення 8. Процес x_t називається **однорідним**, якщо величина $x_{t+h} - x_t$ не залежить від t для будь-якого $t \in \mathbb{T}$ (тобто розподіл такої випадкової величини не залежить від t), а залежить тільки від h і така різниця розподілена точно так само, як і x_h .

§2. Процес Пуассона

Означення. Випадковий процес з неперервним часом і фазовим простором $\{0, 1, 2, \dots\}$ називається **процесом Пуассона**, якщо x_t є процесом з незалежними приростами, однорідним, $x_0 = 0$ і:

1. $P\{x_h = 0\} = 1 - \lambda h + o(h)$ при $h \rightarrow 0$,
2. $P\{x_h = 1\} = \lambda h + o(h)$ при $h \rightarrow 0$,
3. $P\{x_h \geq 2\} = o(h)$ в процесі при $h \rightarrow 0$.

Число λ — називається **параметром процесу**, $0 < \lambda < \infty$, x_t — фактично кількість стрибків процесу до моменту часу t .



Твердження. Якщо x_t — процес Пуассона, то

$$P\{\omega: x_t = k\} = \frac{(\lambda \cdot t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \dots$$

◀ Для доведення використовується апарат характеристичних функцій. Зафіксуємо довільне $t > 0$. Позначимо характеристичну функцію випадкової величини x_t (при фіксованому t це випадкова величина) $\psi_t(u)$, де $u \in \mathbb{R}$. За означенням характеристичної функції,

$$\psi_{t+h}(u) = \mathbb{E} e^{i x_{t+h} u} = \mathbb{E} e^{i [x_{t+h} - x_t + x_t] u} =$$

оскільки процес є процесом з незалежними приростами, то випадкові величини x_t і $x_{t+h} - x_t$ — незалежні, а тому і їх

експоненти незалежні, а звідси математичне сподівання добутку дорівнює добутку математичних сподівань:

$$= \mathbb{E} \left(e^{i[x_{t+h}-x_t]u} \cdot e^{ix_t u} \right) = \mathbb{E} e^{i[x_{t+h}-x_t]u} \cdot \mathbb{E} e^{ix_t u} =$$

$= \psi_t(u)$

але процес однорідний, тому приріст $x_{t+h} - x_t$ не залежить від t , а лише від h , а отже,

$$= \mathbb{E} e^{ix_h u} \cdot \psi_t(u).$$

Про x_h відомо таке:

інші значення

x_h	0	1	A
P	$1 - \lambda \cdot h + o(h)$	$\lambda \cdot h + o(h)$	$o(h)$

і тоді математичне сподівання дорівнює:

$$\mathbb{E} e^{ix_h u} = 1 \cdot (1 - \lambda \cdot h + o(h)) + e^{iu} (\lambda \cdot h + o(h)) + K \cdot o(h),$$

тут $K = e^{iAu} \cdot o(h)$ – стала, фактично

$$\mathbb{E} e^{ix_h u} = 1 - \lambda \cdot h (1 - e^{iu}) + o(h),$$

тоді одержуємо таке:

$$\psi_{t+h}(u) = \psi_t(u) [1 - \lambda h (1 - e^{iu})] + \psi_t(z) \cdot o(h)$$

або

$$\psi_{t+h}(u) = \psi_t(u) [1 + \lambda h (e^{iu} - 1)] + \psi_t(u) \cdot o(h).$$

Звідси

$$\frac{\psi_{t+h}(u) - \psi_t(u)}{h} = \lambda (e^{iu} - 1) \cdot \psi_t(u) + \frac{\psi_t(u) \cdot o(h)}{h},$$

але $\frac{o(h)}{h} \rightarrow 0$ при $h \rightarrow 0$, тому

$$\frac{d\psi_t(u)}{dt} = \lambda(e^{iu} - 1) \cdot \psi_t(u)$$

– одержали диференціальне рівняння з такими початковими умовами:

$$\psi_0(u) = 1.$$

Звідси визначається характеристична функція

$$\psi_t(u) = e^{\lambda t(e^{iu} - 1)},$$

а це і є характеристичною функцією випадкової величини (при фіксованому t), розподіленої за законом Пуассона з параметром λt . Тому

$$P\{\omega: x_t = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \dots \quad \blacktriangleright$$

Нехай зафіксовані моменти часу $t_1 < t_2 < \dots < t_n$ для будь-якого n . Спробуємо підрахувати

$$P\{\omega: x_{t_1} = k_1, x_{t_2} = k_2, \dots, x_{t_n} = k_n\} =$$

природно, що $k_1 \leq k_2 \leq \dots \leq k_n$ (бо це кількість стрибків у відповідні моменти часу, і при збільшенні часу, кількість стрибків зростає). Тому остання подія еквівалентна такій:

$$= P\{\omega: x_{t_1} = k_1, x_{t_2} - x_{t_1} = k_2 - k_1, x_{t_3} - x_{t_2} = k_3 - k_2, \dots, x_{t_n} - x_{t_{n-1}} = k_n - k_{n-1}\} =$$

але процес є процесом з незалежними приростами, тому відповідні події незалежні і ймовірність добутку цих подій дорівнює добутку їх імовірностей:

$$= P\{\omega: x_{t_1} = k_1\} \cdot P\{\omega: x_{t_2} - x_{t_1} = k_2 - k_1\} \cdot \dots \cdot P\{\omega: x_{t_n} - x_{t_{n-1}} = k_n - k_{n-1}\} =$$

але процес однорідний, тому ці прирости залежать лише від приросту часу t :

$$= P\{\omega: x_{t_1} = k_1\} \cdot P\{\omega: x_{t_2-t_1} = k_2 - k_1\} \cdot \dots \cdot P\{\omega: x_{t_n-t_{n-1}} = k_n - k_{n-1}\} =$$

А тепер використаємо $P\{\omega: x_t = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$:

$$= \frac{(\lambda t_1)^{k_1}}{(k_1)!} e^{-\lambda t_1} \cdot \frac{[\lambda(t_2 - t_1)]^{k_2 - k_1}}{(k_2 - k_1)!} e^{-\lambda(t_2 - t_1)} \cdot \dots \cdot \frac{[\lambda(t_n - t_{n-1})]^{k_n - k_{n-1}}}{(k_n - k_{n-1})!} e^{-\lambda(t_n - t_{n-1})}.$$

Тобто сумісний розподіл повністю задається і процес можна вивчати у вузькому розумінні.

У літературі позначають

$$\gamma(t) = \frac{x_t(\omega) - m_{x_t}(t)}{\sigma_{x_t}(t)}$$

– **стандартизований пуасонівський процес.**

Це був один приклад випадкового процесу. Тепер наведемо інший приклад.

§3. Процес Вінера* (вінерівський процес)

Означення. Процес x_t називається **процесом Вінера**, якщо це неперервний за часом процес, з неперервним фазовим простором (фазовим простором $\mathbb{R}$), який є процесом з незалежними приростами, однорідним, $x_0(\omega) = 0$ і

1. $\mathbb{E}x_h = ah + o(h)$ при $h \rightarrow 0$, де $a \in \mathbb{R}$,
2. $\mathbb{E}x_h^2 = bh + o(h)$ при $h \rightarrow 0$, де $0 < b < +\infty$,
3. $\mathbb{E}|x_h|^3 = o(h)$ – третій абсолютний момент.

В означенні фігурують параметри, що визначають процес: a – називається коефіцієнтом переносу, b – коефіцієнтом дифузії.

Твердження. Якщо x_t процес Вінера, то

$$P\{\omega: x_t < x\} = \frac{1}{\sqrt{2\pi bt}} \int_{-\infty}^x e^{-\frac{(u-at)^2}{2bt}} du, \quad x \in \mathbb{R}.$$

Зліва записана функція розподілу $F_{x_t}(x)$, а справа – функція розподілу нормально розподіленої випадкової величини $N(at, bt)$, тобто у твердженні, фактично, постулюється таке: при кожному t маємо нормальний розподіл.

◀ Для доведення застосовується підхід, як і в попередньому параграфі. Знайдемо характеристичну функцію: $\psi_t(u) = \mathbb{E}e^{iux_t}$.

Спочатку випишемо приріст:

$$\psi_{t+h}(u) = \mathbb{E}e^{iux_{t+h}} = \mathbb{E}\left(e^{iu[x_{t+h}-x_t]} \cdot e^{iux_t}\right) =$$

використовуємо незалежність приростів і однорідність

* Норберт Вінер (Norbert Wiener, 1894 – 1964) – американський математик.

$$= \mathbb{E} e^{iux_h} \cdot \psi_t(u) =$$

враховуючи умови твердження, розкладемо e^{ix} в ряд Маклорена^{*}

$$\begin{aligned} &= \mathbb{E} \left[1 + iux_h - \frac{1}{2}u^2 x_h^2 - \frac{1}{3!}iu^3 x_h^3 \right] \cdot \psi_t(u) = \\ &= \left(1 + iahu - \frac{1}{2}bhu^2 + o(h) \right) \cdot \psi_t(u). \end{aligned}$$

Тепер треба знайти приріст і поділити на h :

$$\frac{\psi_{t+h}(u) - \psi_t(u)}{h} = \left(iau - \frac{1}{2}bu^2 \right) \cdot \psi_t(u) + \psi_t(u) \cdot \frac{o(h)}{h}$$

тут $\frac{o(h)}{h} \rightarrow 0$, при $h \rightarrow 0$, тоді одержимо диференціальне рівняння:

$$\frac{d\psi_t(u)}{dt} = \left(iau - \frac{1}{2}bu^2 \right) \cdot \psi_t(u)$$

з початковою умовою:

$$\psi_0(u) = 1.$$

Розв'язок цієї задачі при довільному t записується так:

$$\psi_t(u) = e^{iu(at) - \frac{u^2(\sqrt{bt})^2}{2}},$$

тобто

$$\psi_t(u) = e^{iuat - \frac{u^2 bt}{2}},$$

а це і є характеристична функція випадкової величини, розподіленої нормально з параметрами $N(at, bt)$.

^{*} К'олін Макл'орен (Colin Maclaurin, 1698-1746) – шотландський математик.

Отже, при кожному t випадкова величина x_t розподілена нормально з математичним сподіванням at і середнім квадратичним відхиленням $\sqrt{bt}$: $N(at, bt)$. Таким чином, постульований факт доведено: дійсно, процес Вінера при кожному t має нормальний закон розподілу. ►

Як правило, розглядається стандартизований процес Вінера і його позначають

$$w(t) = \frac{x_t - at}{\sqrt{bt}}.$$

Для нього $E(w(t)) = 0$, $\sigma_{w(t)} = 1$.

Аналогічно, як і для процесу Пуассона, встановлюється закон сумісного розподілу для будь-якого набору $t_1, t_2, \dots, t_n$ у вигляді функції сумісного розподілу:

$$P\{\omega: x_{t_1} < x_1, x_{t_2} < x_2, \dots, x_{t_n} < x_n\},$$

тут $x_1, x_2, \dots, x_n \in \mathbb{R}$. Одержується багатовимірний нормальний розподіл.

§4. Стохастичний інтеграл Іто*.

Стохастичне диференціальне рівняння

Будемо вважати, що $a = 0$, $b = 1$ в означенні стохастичного вінерівського процесу. Тобто розглядається процес $w(t)$, для якого

$$\mathbb{E}w(t) = o(h), \quad \mathbb{E}w^2(t) = h + o(h), \quad \mathbb{E}|w(t)|^3 = o(h).$$

Цей процес $w(t)$ володіє такими властивостями:

- з імовірністю одиниця процес $w(t)$ неперервний при кожному t ;
- з імовірністю одиниця процес $w(t)$ не є функцією з обмеженою варіацією;
- з імовірністю одиниця процес $w(t)$ не є диференційовною функцією при будь-якому t .

Побудуємо такі конструкції: процес будемо розглядати на скінченному відрізку часу $[0, T]$; будемо вважати, що задані такі σ -алгебри подій $\mathcal{F}_t$, що

- 1) із $t_1 < t_2$ випливає $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$,
- 2) процес $w(t)$ – вимірний відносно $\mathcal{F}_t$,
- 3) приріст $w(t+s) - w(t)$ не залежить від $\mathcal{F}_t$ при довільних фіксованих $t, s > 0$. Тоді кажуть, що вінерівський процес $w(t)$ узгоджений з потоком σ -алгебр.

Позначимо також через $H_2[0, T]$ простір вимірних, простих і сумовних у середньому квадратичному функцій, тобто простір (множину) всіх випадкових процесів, вимірних відносно $\mathcal{F}_t$, таких, що з імовірністю одиниця існує

* *Kiyosi Ito (伊藤 清, 1915 – 2008) – японський математик.*

$$\int_0^T f^2(t) dt$$

для $f(t) \in H_2[0, T]$, тобто

$$P \left\{ \int_0^T f^2(t) dt < \infty \right\} = 1.$$

Для таких $f(t)$ можна ввести в розгляд стохастичний інтеграл за вінерівським процесом

$$\int_0^T f(t) dw(t)$$

так:

$$\int_0^T f(t) dw(t) = p \lim_{\substack{\Delta \rightarrow 0 \\ (n \rightarrow \infty)}} \sum_{k=0}^{n-1} f(t_k) [w(t_{k+1}) - w(t_k)],$$

де $t_0, t_1, t_2, \dots, t_n$ – розбиття відрізка $[0, T]$, $0 = t_0 < t_1 < t_2 < \dots < t_n = T$,
 $\Delta = \max_k (t_{k+1} - t_k)$.

Можна довести, що такий інтеграл має традиційні властивості:

1) якщо $f(t) \in H_2[0, T]$, то інтеграл $\int_a^b f(t) dw(t)$ існує для довільних

$$a, b \in [0, T];$$

2) $\int_0^t f(t) dw(t)$ вимірний відносно σ -алгебри $\mathcal{F}_t$;

3) адитивність $\int_a^b [cf_1(t) + df_2(t)]dw(t) = c \int_a^b f_1(t)dw(t) + d \int_a^b f_2(t)dw(t);$

4) математичне сподівання інтеграла дорівнює нулю:

$$\mathbb{E} \int_a^b f(t)dw(t) = 0,$$

5) математичне сподівання квадрата інтеграла дорівнює:

$$\mathbb{E} \left(\int_a^b f(t)dw(t) \right)^2 = \int_a^b \mathbb{E} f^2(t)dt.$$

Тут у правій частині рівності записаний звичайний інтеграл Рімана.

Розглянемо приклади:

Приклад 1. $\int_a^b dw(t) = w(b) - w(a).$

Приклад 2. $\int_a^b w(t)dw(t).$ Випишемо інтегральну суму:

$$\sum_{k=0}^{n-1} w(t_k)[w(t_{k+1}) - w(t_k)] = \sum_{k=0}^{n-1} [w(t_k)w(t_{k+1}) - w^2(t_k)] =$$

| виділимо повний квадрат |

$$= \frac{1}{2} \sum_{k=0}^{n-1} (w^2(t_{k+1}) - w^2(t_k)) - \frac{1}{2} \sum_{k=0}^{n-1} (w(t_{k+1}) - w(t_k))^2$$

перейдемо до границі при $\Delta \rightarrow 0$ ($n \rightarrow \infty$) за ймовірністю і одержимо:

$$\int_a^b w(t)dw(t) = \frac{1}{2}w^2(b) - \frac{1}{2}w^2(a) - \frac{1}{2}(b-a).$$

Цей інтеграл називається **стохастичним інтегралом Вінера – Іто**.

Нехай розглядається детермінована функція $y(t)$ на $[0, T]$ (не випадковий процес). Припустимо, що $y(t)$ – диференційовна функція. Тоді диференціал $dy(t) = y'(t)dt$. Для будь-яких t_1, t_2 можна записати

$$y(t_2) - y(t_1) = \int_{t_1}^{t_2} y'(t)dt.$$

Пригадаємо: якщо $y(t)$ така функція, що для будь-яких t_1, t_2 з відрізка $[0, T]$ можна записати

$$y(t_2) - y(t_1) = \int_{t_1}^{t_2} a(t)dt,$$

то $a(t)dt$ – диференціал функції $y(t)$.

Припустимо, що існують такі випадкові процеси (функції) $a(t)$, $b(t)$, що $\sqrt{|a(t)|} \in H_2[0, T]$, $b(t) \in H_2[0, T]$ і для будь-яких $t_1, t_2 \in [0, T]$ можна записати

$$x(t_2) - x(t_1) = \int_{t_1}^{t_2} a(t)dt + \int_{t_1}^{t_2} b(t)dw(t).$$

Тоді кажуть, що існує **стохастичний диференціал** процесу $x(t)$ і записують

$$dx(t) = a(t)dt + b(t)dw(t).$$

Відмінності від звичайного диференціала:

якщо для $x_1(t)$ стохастичний диференціал

$$dx_1(t) = a_1(t)dt + b_1(t)dw(t),$$

для $x_2(t)$ стохастичний диференціал

$$dx_2(t) = a_2(t)dt + b_2(t)dw(t),$$

то стохастичний диференціал добутку дорівнює

$$d(x_1(t)x_2(t)) = x_1(t)dx_2(t) + x_2(t)dx_1(t) + \frac{1}{2}b_1(t)b_2(t)dw(t).$$

Диференціал суми дорівнює сумі диференціалів.

Якщо $x(t)$ має стохастичний диференціал

$$dx(t) = a(t)dt + b(t)dw(t),$$

а функція $f(t, x)$ – неперервна і має неперервні похідні $f'_t(t, x)$, $f'_x(t, x)$, $f''_{xx}(t, x)$, то стохастичний диференціал для $f(t, x(t))$ (диференціал складеної функції) такий:

$$df(t, x(t)) = \left[f'_t(t, x(t)) + \frac{1}{2} f''_{xx}(t, x) b^2(t) + f'_x(t, x(t)) a(t) \right] dt + f'_x(t, x(t)) b(t) dw(t).$$

Ця формула називається **формулою Іто**.

Можна розглядати і більш складні структури стохастичних диференціалів.

Якщо введено поняття стохастичного диференціала, то можна ввести в розгляд поняття стохастичного диференціального рівняння.

У звичайних диференціальних рівняннях розглядали рівняння $y'(t) = f(t, y(t))$. Диференціал функції, що є розв'язком цього рівняння $dy(t) = f(t, y(t))dt$. Можна було писати й інакше

$$y(t_2) - y(t_1) = \int_{t_1}^{t_2} f(t, y(t))dt.$$

Аналогічно введемо поняття у стохастичному випадку. Якщо для будь-яких t_1, t_2 ($t_1 < t_2$), процес можна представити так

$$x(t_2) - x(t_1) = \int_{t_1}^{t_2} a(t, x(t))dt + \int_{t_1}^{t_2} b(t, x(t))d\omega(t),$$

то кажуть, що процес $x(t)$ є розв'язком стохастичного диференціального рівняння

$$dx(t) = a(t, x(t))dt + b(t, x(t))d\omega(t).$$

Останнє рівняння – це **стохастичне диференціальне рівняння Іто**. Його розв'язок – дифузійний процес із коефіцієнтом дифузії $b^2(t, x)$ і коефіцієнтом переносу $a(x, t)$

Це рівняння стає класичним, якщо $b(t, x(t)) = 0$ і всі результати одержані для стохастичних рівнянь повинні узагальнювати класичні результати.

При допомозі другого доданка (випадкового) можна моделювати вплив на динамічну систему випадкових факторів неперервного типу.

Виявляється, що, використовуючи відомий процес Пуассона, можна побудувати більш загальне стохастичне диференціальне рівняння, в якому враховуються дії на систему випадкових дискретних збурень (або поштовхів).

Аналогічно до попереднього можна ввести в розгляд стохастичний диференціал за процесами Вінера і Пуассона. Такий стохастичний диференціал має вигляд

$$dx(t) = a(t)dt + b(t)dw(t) + \int_{\mathbb{R}} \Gamma(t, u) \tilde{\nu}(dt, du),$$

де $\tilde{\nu}(t, A) = \nu(t, A) - \Pi(A)(t - t_0)$, $\nu(t, A)$ – кількість стрибків процесу Пуассона, величини яких попадають в борелівську множину A , побудовану на додатній півосі, $\Pi(A)$ – міра цієї множини.

Відповідне диференціальне рівняння має вигляд

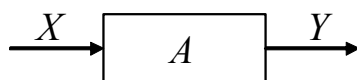
$$dx(t) = a(t, x(t))dt + b(t, x(t))dw(t) + \int_{\mathbb{R}} \Gamma(t, u, x(t)) \tilde{\nu}(dt, du)$$

з урахуванням випадкових поштовхів.

§5. Канонічне представлення стаціонарного випадкового процесу

На динамічні системи, особливо на системи автоматичного керування, випадкові фактори часто мають істотний вплив. Тому важливе питання про те, як перетворюються випадкові сигнали, які діють на такі системи.

Розглядатимемо приклади найпростіших лінійних систем. Можна уявити собі таке: на вхід системи A надходить сигнал X і на виході маємо Y . Сигнал X включає випадкові збурення, які є адитивною складовою сигналу.



Якщо розглядати лінійні системи, то

$$A[x + \xi] = A[x] + A[\xi].$$

Тому надалі будемо вважати, що на вхід системи надходить випадковий процес, який будемо позначати $X(t)$, на виході системи утворюється випадковий процес $Y(t)$.

За характеристиками випадкового процесу $X(t)$ необхідно встановити характеристики випадкового процесу $Y(t)$. Надалі будемо вести мову про дослідження процесу в широкому розумінні, тобто нас будуть цікавити лише математичні сподівання процесу і автокореляційна функція. Вивчити процес у вузькому розумінні дуже складно. Отже, надалі нас будуть цікавити такі характеристики для випадкового процесу $X(t)$: математичне сподівання

$$\mathbb{E}X(t) = m_x(t)$$

та автокореляційна функція

$$K_x(t, t') = \mathbb{E}[(X(t) - m_x(t))(X(t') - m_x(t'))]$$

і для випадкового процесу $Y(t)$ також

$$\mathbb{E}Y(t) = m_y(t) \quad \text{та} \quad K_y(t, t').$$

Фактично вважатимемо, що відомі $m_x(t)$ і $K_x(t, t')$. Як правило вважається, що $t', t \in [0, T]$ або й $[0, +\infty)$. Для лінійних систем поставлена задача знаходження $m_y(t)$ і $K_y(t, t')$ розв'язується досить просто.

Розглянемо найпростіші системи.

Приклад 1. $Y(t) = a \cdot X(t)$, де $a \in \mathbb{R}$.

Нехай відомо $m_x(t)$ і $K_x(t, t')$. За властивостями математичного сподівання, матимемо

$$m_y(t) = \mathbb{E}[a \cdot X(t)] = a \cdot m_x(t),$$

тобто математичне сподівання перетворюється точно за тим самим законом, за яким перетворюється процес $X(t)$. Розглянемо

$$\begin{aligned} K_y(t, t') &= \mathbb{E}[(Y(t) - m_y(t)) \cdot (Y(t') - m_y(t'))] = \\ &= \mathbb{E}[a(X(t) - m_x(t)) \cdot a(X(t') - m_x(t'))] = a \cdot a \cdot K_x(t, t'), \end{aligned}$$

тобто маємо таку ситуацію:

$$m_y(t) = a \cdot m_x(t), \quad K_y(t, t') = a \cdot (a \cdot K_x(t, t')),$$

тобто автокореляційна функція перетворюється за тим самим законом, що й процес $X(t)$, двічі.

Приклад 2. Задана випадкова функція $X(t)$ з математичним сподіванням $m_x(t)$ і кореляційною функцією $K_x(t, t')$. Випадкова

функція $Y(t)$ зв'язана з $X(t)$ лінійним однорідним оператором інтегрування:

$$Y(t) = \int_0^t X(\tau) d\tau.$$

Потрібно знайти характеристики випадкової функції $Y(t)$: $m_y(t)$ і $K_y(t, t')$. Знову одержимо

$$m_y(t) = \int_0^t m_x(\tau) d\tau,$$

тобто математичне сподівання інтеграла від випадкової функції дорівнює інтегралу від її математичного сподівання. Іншими словами, операцію інтегрування і операцію математичного сподівання можна поміняти місцями.

Знайдемо кореляційну функцію $K_y(t, t')$. Для цього перейдемо до центрованих випадкових величин:

$$\begin{aligned} K_y(t, t') &= \mathbb{E} \left[\overset{\circ}{Y}(t) \cdot \overset{\circ}{Y}(t') \right] = \mathbb{E} \left(\int_0^t \overset{\circ}{X}(\tau) d\tau \cdot \int_0^{t'} \overset{\circ}{X}(u) du \right) = \\ &= \mathbb{E} \int_0^t \int_0^{t'} \overset{\circ}{X}(\tau) \overset{\circ}{X}(u) d\tau du = \int_0^t \int_0^{t'} \mathbb{E} \left[\overset{\circ}{X}(\tau) \cdot \overset{\circ}{X}(u) \right] d\tau du = \int_0^t \int_0^{t'} K_x(\tau, u) d\tau du. \end{aligned}$$

Таким чином, для того щоб знайти кореляційну функцію від інтеграла від випадкової функції, потрібно двічі проінтегрувати кореляційну функцію вихідної випадкової функції – спочатку за одним аргументом, потім – за другим:

$$K_y(t, t') = \int_0^t \left(\int_0^{t'} K_x(\tau, u) du \right) d\tau.$$

Іншими словами, кореляційна функція перетворюється двічі за тим самим законом, що й сам випадковий процес $X(t)$.

Приклад 3. Задана випадкова функція $X(t)$ з математичним сподіванням $m_x(t)$ і кореляційною функцією $K_x(t, t')$. Випадкова функція $Y(t)$ зв'язана з випадковою функцією $X(t)$ лінійним однорідним оператором диференціювання

$$Y(t) = \frac{d}{dt} X(t).$$

Тоді $m_y(t) = \frac{d}{dt} m_x(t)$, тобто математичне сподівання похідної від випадкової функції дорівнює похідній від її математичного сподівання. Отже, операцію диференціювання, як і операцію інтегрування, також можна поміняти місцями з операцією математичного сподівання.

Крім цього,

$$K_y(t, t') = \frac{\partial}{\partial t} \left(\frac{\partial}{\partial t'} K_x(t, t') \right).$$

Таким чином, щоб знайти кореляційну функцію похідної, потрібно двічі продиференціювати кореляційну функцію вихідної випадкової функції – спочатку за одним аргументом, потім – за другим, тобто знову той самий оператор застосовується двічі.

І взагалі, якщо випадкова функція $X(t)$ з математичним сподіванням $m_x(t)$ і кореляційною функцією $K_x(t, t')$ перетворюється у випадкову функцію $Y(t)$ лінійним однорідним оператором A (тобто система лінійна): $Y(t) = A[X(t)]$, то для знаходження математичного сподівання випадкової функції $Y(t)$ потрібно застосувати той самий оператор до математичного сподівання випадкової функції $X(t)$:

$$m_y(t) = A[m_x(t)].$$

А для знаходження кореляційної функції потрібно двічі застосувати той самий оператор до кореляційної функції випадкової функції $X(t)$ – спочатку за одним аргументом, а потім – за другим:

$$K_y(t, t') = A^{(t)} [A^{(t')} [K_x(t, t')]],$$

тут A – лінійний оператор, $A^{(t)}$ – цей самий оператор за аргументом t , $A^{(t')}$ – за аргументом t' .

Аналогічні прості результати можна отримати, якщо $U(t)$ є сумою

$$U(t) = X(t) + Y(t)$$

двох відомих процесів $X(t)$, $Y(t)$. У цьому випадку

$$m_u(t) = A[m_x(t)] + A[m_y(t)].$$

Для визначення кореляційної функції $K_u(t, t')$ перейдемо до центрованих випадкових величин:

$$\begin{aligned} K_u(t, t') &= \mathbb{E}[\overset{\circ}{U}(t) \cdot \overset{\circ}{U}(t')] = \\ &= \mathbb{E}[(X + Y - m_x - m_y)(t) \times (X - Y - m_x - m_y)(t')] = \\ &= \mathbb{E}[(X - m_x)_t (X - m_x)_{t'}] + \mathbb{E}[(Y - m_y)_t (Y - m_y)_{t'}] + \\ &\quad + \mathbb{E}[(X - m_x)_t (Y - m_y)_{t'}] + \mathbb{E}[(X - m_x)_{t'} (Y - m_y)_t] = \\ &= K_x(t, t') + K_y(t, t') + K_{xy}(t, t') + K_{xy}(t', t). \end{aligned}$$

Тут останні два доданки відображають взаємну кореляцію доданків X і Y . Тобто автокореляційна функція для вихідного сигналу $U(t)$ дорівнює сумі автокореляційних функцій для кожного з доданків і ще сумі взаємно кореляційних функцій цих доданків.

У випадку, якщо вхідні (адитивні) сигнали некорельовані (а ще краще незалежні), то $K_{xy} = 0$ і отримаємо

$$K_u(t, t') = K_x(t, t') + K_y(t, t').$$

Розглянемо найпростіші випадки вхідних сигналів. Припустимо, що вхідний сигнал має вигляд

$$X(t) = U\varphi(t),$$

де U – випадкова величина (скалярна), а $\varphi(t)$ – не випадкова функція. Будемо вважати, що математичне сподівання дорівнює нулю (до такої випадкової величини завжди можна перейти):

$$\mathbb{E}U = 0, \quad m_x(t) = 0, \quad \varphi(t) \neq 0.$$

Тоді

$$K_x(t, t') = \mathbb{E}[U\varphi(t) \cdot U\varphi(t')] = \mathbb{E}U^2 \cdot \varphi(t) \cdot \varphi(t') = D_u \varphi(t) \varphi(t').$$

Сигнал $X(t)$ такого виду називають **найпростішим випадковим сигналом**. Дисперсія $DX(t)$ у цьому випадку така:

$$DX(t) = K_x(t, t) = D_u [\varphi(t)]^2.$$

Нехай сигнал трохи складніший:

$$X(t) = \sum_{k=1}^n U_k \varphi_k(t), \quad \mathbb{E}U_k = 0, \quad (k = \overline{1, n}).$$

Знайдемо кореляційну функцію і дисперсію випадкової функції $X(t)$:

$$K_x(t, t') = \sum_{k=1}^n D_k \varphi_k(t) \varphi_k(t') + \sum_{\substack{i=1 \\ (i \neq j)}}^n \sum_{j=1}^n \mathbb{E}[U_i U_j] \cdot \varphi_i(t) \varphi_j(t'),$$

де D_k – дисперсія випадкової величини U_k . Другий доданок тут дорівнює нулю, якщо U_i та U_j ($i \neq j$) некорельовані (якщо незалежні, то тим більше). Такі випадкові процеси

$$X(t) = \sum_{k=1}^n U_k \varphi_k(t), \quad (1)$$

для яких коефіцієнти U_i та U_j ($i \neq j$) некорельовані, називаються **канонічними випадковими процесами**.

При цьому функції $\varphi_k(t)$ називаються **координатними функціями**, а випадкові величини U_k – **амплітудами процесу**, що відповідають відповідним координатним функціям.

Для канонічних випадкових процесів кореляційна функція

$$K_x(t, t') = \sum_{k=1}^n D_k \varphi_k(t) \varphi_k(t'), \quad (2)$$

де D_k – дисперсії амплітуд U_k процесу. Вираз (2) називається **канонічним розкладом кореляційної функції**.

Отже, для канонічного випадкового процесу (1) можна знайти канонічний розклад його кореляційної функції (2). Можна довести, що при деяких досить слабких додаткових умовах на функції $\varphi_k(t)$ обернене твердження також правильне, а саме: якщо канонічний розклад автокореляційної функції процесу має вигляд (2), то випадковий процес $X(t)$ може бути представлений у вигляді (1).

Цей результат належить Пугачову* і таке представлення (1) називається представленням Пугачова, де U_k – деякі випадкові величини з математичними сподіваннями, що дорівнюють нулю.

Аналогічні результати отримані Пугачовим для нескінченних сум, а також для інтегральних представлень процесу.

* Володимир Семенович Пугачов (1911-1998) – радянський математик.

Нехай вхідний процес $X(t)$ заданий автокореляційною функцією $K_x(t, t')$. Припустимо, що це стаціонарний процес (сумісні розподіли залишаються тими самими при зміщенні по часу). У цьому випадку $K_x(t, t')$ залежить лише від $t - t'$ і є парною функцією:

$$K_x(t, t') = k(t - t') = k(t' - t),$$

тобто $k(\tau)$ де $\tau = t' - t$ і $k(\tau)$ – парна функція, якщо процес $X(t)$ заданий на $t \in [0, T]$, то $-T \leq \tau \leq T$.

Припустимо, що $k(\tau)$ задана. Таку функцію на $[0, T]$ можна представити наближено з достатньою точністю відрізком ряду Фур'є, використовуючи тільки парні (косинусні) гармоніки, у такому вигляді:

$$\begin{aligned} k(\tau) &\approx \sum_{k=0}^n a_k \cos \frac{k\pi}{T} \tau = \sum_{k=0}^n a_k \cos \frac{k\pi}{T} (t - t') = \\ &= \sum_{k=0}^n a_k \left(\cos \frac{k\pi}{T} t \cos \frac{k\pi}{T} t' + \sin \frac{k\pi}{T} t \sin \frac{k\pi}{T} t' \right) = \\ &= \sum_{k=0}^n a_k \cos \frac{k\pi}{T} t \cos \frac{k\pi}{T} t' + \sum_{k=0}^n a_k \sin \frac{k\pi}{T} t \sin \frac{k\pi}{T} t', \end{aligned}$$

тобто одержали канонічний розклад кореляційної функції $K_x(t, t')$. Координатними функціями цього канонічного розкладу є по чергово косинуси і синуси частот, кратних $\frac{\pi}{T}$.

Вище зазначали, що за канонічним розкладом кореляційної функції можна побудувати канонічний розклад самої випадкової функції з тими ж координатними функціями і з дисперсіями, рівними коефіцієнтам a_k у канонічному розкладі кореляційної функції. Тобто виявляється, що випадковий процес $X(t)$ можна наближено предста-

вити у вигляді

$$X(t) \approx \sum_{k=0}^n U_k \cos \frac{k\pi}{T} t + \sum_{k=0}^n V_k \sin \frac{k\pi}{T} t,$$

де U_k, V_k – некорельовані випадкові величини з математичними сподіваннями, що дорівнюють нулю: $\mathbb{E}U_k = \mathbb{E}V_k = 0$ і дисперсіями, однаковими для кожної пари випадкових величин з одним і тим самим індексом k : $DU_k = DV_k = a_k$.

Взагалі кажучи, якщо представити кореляційну функцію у вигляді ряду

$$k(\tau) = \sum_{k=0}^{\infty} a_k \cos \frac{k\pi}{T} \tau,$$

то випадковий процес $X(t)$ може бути представлений на відрізку $[-T, T]$ у такому вигляді:

$$X(t) = \sum_{k=0}^{\infty} \left(U_k \cos \frac{k\pi}{T} \tau + V_k \sin \frac{k\pi}{T} t \right),$$

де U_k, V_k – випадкові величини з $\mathbb{E}U_k = \mathbb{E}V_k = 0$, $DU_k = DV_k = a_k$.

Набір чисел $a_k, k = \overline{0, \infty}$ називається **спектром випадкового процесу**.

Легко зрозуміти, що дисперсія стаціонарного випадкового процесу (при такому представленні) буде дорівнювати сумі дисперсій всіх гармонік його спектрального розкладу:

$$D_x = \sum_{k=0}^{\infty} \left(\cos^2 \frac{k\pi}{T} \tau + \sin^2 \frac{k\pi}{T} t \right) a_k^2 = \sum_{k=0}^{\infty} a_k^2.$$

При $T \rightarrow \infty$ відповідні ряди Фур'є перетворюються в інтеграли Фур'є і тоді

$$k(\tau) = \int_0^{\infty} S(\omega) \cos \omega \tau d\omega,$$

де

$$S(\omega) = \frac{2}{\pi} \int_0^{\infty} k(\tau) \cos \omega \tau d\tau,$$

функція $S(\omega)$ називається **спектральною щільністю випадкового процесу**.

Отже кореляційна функція і спектральна щільність виражаються одна через одну за допомогою перетворення Фур'є. Часто використовують нормовану спектральну щільність

$$\frac{S(\omega)}{D_x}.$$

За допомогою аналізу спектра випадкового процесу або спектральної щільності вивчається в широкому розумінні поведінка процесу на скінченному або нескінченному інтервалі.

§6. Поняття про марковський процес (процес Маркова)

Розглядатимемо процес з дискретним часом і дискретним фазовим простором; тобто t набуває значення: $t = 0, 1, 2, \dots$

Фазовий простір X (множина значень випадкового процесу) – скінченна множина. Випадковий процес $x(t)$ ($x_t(\omega)$) може набувати у будь-який момент часу $t = 0, 1, 2, \dots$ будь-які значення з X .

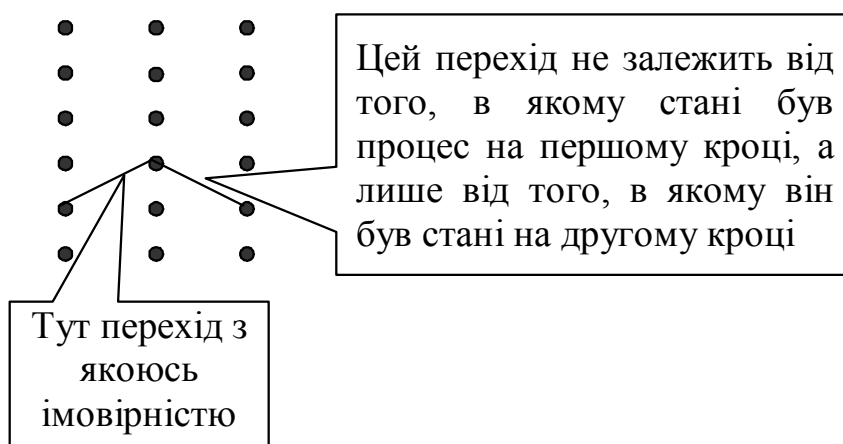
Маємо фактично справу з послідовністю випадкових величин. Такий процес називається **ланцюгом**.

Означення. Описаний дискретний випадковий процес називається **ланцюгом Маркова**, якщо ймовірність того, що

$$\begin{aligned} P\{x(n) = y \mid x(0) = x_0, x(1) = x_1, \dots, x(n-1) = x\} = \\ = P\{x(n) = y \mid x(n-1) = x\}, \end{aligned}$$

де $x, y, x_1, \dots, x_{n-2} \in X$.

Тобто ймовірність того, що $x(n) = y$ не залежить від всієї історії розвитку, а лише від того, в якому стані знаходиться процес перед цим.



Означення. Ймовірність

$$P\{x(n)=y \mid x(n-1)=x\}=P_n(x,y),$$

якщо вона існує, називається **перехідною функцією** марківського ланцюга (процесу).

Зрозуміло, що $P_n(x,y) \geq 0$. Якщо $P_n(x,y)=0$, то можна довізнати перехідну функцію умовою $\sum_{y \in X} P_n(x,y)=1$.

Тому більш загально перехідну функцію визначають так: це така невід'ємна функція $P_n(x,y) \geq 0$, що $\sum_{y \in X} P_n(x,y)=1$ і ця функція

$P_n(x,y)$ дорівнює

$$P\{x(n)=y \mid x(n-1)=x\}=P_n(x,y)$$

при умові, що така ймовірність існує. Часто таку ймовірність називають **ймовірністю переходу з одним кроком**.

Зауважимо: якщо ця ймовірність не залежить від n , тобто $P_n(x,y)=P(x,y)$, то маємо стаціонарний марківський процес.

Властивості функції переходу

Обчислимо (підрахуємо)

$$P\{x(0)=x_0, x(1)=x_1, x(2)=x_2, \dots, x(k)=x_k\}=$$

| за теоремою множення ймовірностей |

$$= P\{x(0)=x_0\} \cdot P\{x(1)=x_1 \mid x(0)=x_0\} \times$$

$$\times P\{x(2)=x_2 \mid x(0)=x_0, x(1)=x_1\} \times \dots \times$$

$$\times P\{x(k)=x_k \mid x(0)=x_0, x(1)=x_1, \dots, x(k-1)=x_{k-1}\} =$$

за означенням марківського процесу:

$$= P\{x(0) = x_0\} \cdot P_1(x_0, x_1) \cdot P_2(x_1, x_2) \cdot \dots \cdot P_k(x_{k-1}, x_k).$$

Тобто, якщо відома перехідна функція, як функція часу(індекса), то можна виписати всі можливі сумісні розподіли. Отже, марківський процес повністю задається функцією $P_k(x, y)$ від x, y і часу k .

Введемо таке позначення: вважаємо, що деяке m більше за n ($m > n$) і

$$P_{n,m}(x, y) = \sum_{x_{n+1} \in X} \sum_{x_{n+2} \in X} \dots \sum_{x_{m-1} \in X} P_{n+1}(x, x_{n+1}) \cdot P_{n+2}(x_{n+1}, x_{n+2}) \cdot \dots \cdot P_m(x_{m-1}, y),$$

тут суми по всіх можливих переходах.

Оскільки розглядають скінченний фазовий простір, то всі суми скінченні; якщо б розглядати злічений фазовий простір, то були б ряди.

Зауважимо, що $P_{n,n+1}(x, y) = P_{n+1}(x, y)$.

Тепер формально отримаємо те, що інтуїтивно цілком зрозуміло: $P_{n,m}(x, y)$ – це функція переходу з x в y за $m - n$ кроків.

Випишемо

$$\begin{aligned} P\{x(n) = x, x(n+1) = x_{n+1}, x(n+2) = x_{n+2}, \dots, x(m) = y\} = \\ = P\{x(n) = x\} \cdot P_{n+1}\{x, x_{n+1}\} \cdot P_{n+2}\{x_{n+1}, x_{n+2}\} \cdot \dots \cdot P_m\{x_{m-1}, y\}. \end{aligned}$$

Просумуємо цю рівність за всіма можливими $x_{n+1}, x_{n+2}, \dots, x_{m-1}$ з X . Одержимо:

$$P\{x(n) = x, x(m) = y\} = P\{x(n) = x\} \cdot P_{n,m}(x, y).$$

Звідси

$$P\{x(n) = x\} \cdot P\{x(m) = y \mid x(n) = x\} = P\{x(n) = x\} \cdot P_{n,m}(x, y),$$

тоді

$$P\{x(m) = y \mid x(n) = x\} = P_{n,m}(x, y).$$

З отриманих співвідношень випливає таке співвідношення для ймовірностей переходу за один крок: при $k < n < m$

$$P_{k,m}(x, y) = \sum_{u \in X} P_{k,n}(x, u) \cdot P_{n,m}(u, y)$$

– ця рівність називається **рівнянням Колмогорова – Чепмена***.

Рівняння Колмогорова – Чепмена означає, що ймовірність переходу зі стану x в момент k в стан y в момент m дорівнює сумі по всіх можливих $u \in X$ переходах зі стану x в момент k у стан u в проміжний момент n і зі стану u в момент n у стан y в момент m .

Будемо вважати, що фазовий простір X складається з n точок (станів): $X = \{x_1, \dots, x_n\}$. Крім цього, будемо вважати, що процес стаціонарний (ймовірність переходу за один крок не залежить від часу):

$$P_n\{x(n) = x_j \mid x(n-1) = x_i\} = p_{ij}.$$

Наприклад, p_{12} – ймовірність того, що якщо процес в будь-який момент знаходився в стані x_1 , то в наступний момент часу буде знаходитися в стані x_2 .

Набір ймовірностей p_{ij} задається матрицею Π розміру $N \times N$:

$$\Pi = \left| p_{ij} \right|_{i,j=1}^N.$$

Уведемо у розгляд ймовірності $p_{ij}(k)$ – переходу зі стану x_i в x_j за k кроків.

$$p_{ij}(k) = P\{x(n+k) = x_j \mid x(n) = x_i\}.$$

* Сидні Чепмен (Sydney Chapman, 1888-1970) – англійський математик, фізик і геофізик.

Тоді

$$p_{ij}(n) = \sum_{r=1}^N p_{ir}(s) \cdot p_{rj}(k-s), \quad 0 < s < n.$$

Позначимо $\Pi(1) = \Pi$, $\Pi(2) = |p_{ij}(2)|$. Легко зрозуміти, що $\Pi(2) = \Pi^2$, $\Pi(3) = \Pi^3$ і т. д.

$$\Pi(n) = \Pi^n.$$

Крім цього,

$$\Pi(k) = \Pi(s) \cdot \Pi(k-s).$$

Таким чином, якщо відома матриця переходу за один крок, то можна одержати матрицю переходу за будь-яку кількість кроків. Матриця Π – це ймовірнісна матриця, її елементи $p_{ij} \geq 0$ і сума

$$\sum_{j=1}^N p_{ij} = 1.$$

Виникає питання: як поводить себе матриця Π^n при $n \rightarrow \infty$? Якщо

$$p_{ij}(n) \xrightarrow{n \rightarrow \infty} 1,$$

то стан x_i називається **асимптотично-поглинаючим станом ланцюга Маркова**.

Виявлення поглинаючих станів процесу – одна з найважливіших задач. Якщо для деякого k

$$p_{ii}(k) = 1, \quad p_{ij}(k) = 0, \quad (i \neq j),$$

то момент $t = k$ називається **моментом зупинки процесу**.

Якщо

$$p_{ii}(n) \xrightarrow{n \rightarrow \infty} 0, \quad p_{ji}(n) \xrightarrow[n \rightarrow \infty]{j \neq i} 0$$

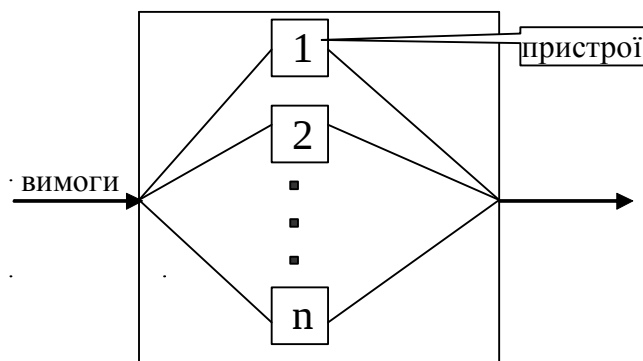
то стан x_i називається **асимптотично відштовхуючим**.

§7. Аналіз випадкових процесів у системах масового обслуговування

Під **системою масового обслуговування** (СМО) розуміють деяку сукупність пристроїв, призначену для виконання (задоволення) вимог, які надходять на вхід цієї сукупності (системи).

Наприклад: СМО – телефонна станція; СМО – перукарня, тоді прилади це перукарі, а вимоги – клієнти.

Якщо система складається з одного обслуговуючого пристрою, то вона називається одноканальною. Відповідно багато пристроїв – багатоканальна. Замість пристрою кажуть ще канал, прилад.



Якщо вимоги надходять у чітко визначені моменти часу і час обслуговування кожної вимоги будь-яким каналом фіксований, то процеси, які відбуваються в системі, називаються регулярними і аналіз такої системи здійснюється відносно просто. Їх розгляд не викликає ніякого інтересу з точки зору теорії ймовірності.

Ситуація ускладнюється, якщо моменти появи вимог випадкові і час обслуговування вимог також має випадковий характер.

Будемо вважати, що всі канали однотипні й усі вимоги однотипні. Тобто вважатимемо, що відомо закон розподілу моментів появи вимог (або у вигляді функції розподілу або у вигляді якихось характеристик). Будемо також вважати, що відомий розподіл (або

характеристики розподілу) довжини часового інтервалу обслуговування вимог кожним каналом.

Усі канали вважатимемо рівноправними (тобто це однотипні прилади).

1. Вхідний потік вимог вважатимемо **найпростішим**, тобто кількість вимог на інтервалі часу $[t_1, t_1 + \Delta t]$ залежить тільки від Δt і не залежить від t_1 — ця вимога називається **стаціонарністю**.

2. Вважатимемо, що потік вимог **ординарний**. Тобто ймовірність того, що кількість вимог $\xi(t)$, які надходять за час Δt , дорівнює нулю, записується так:

$$P\{\xi(t) = 0\} = 1 - \lambda \Delta t + o(\Delta t), \quad \lambda > 0$$

а також

$$P\{\xi(t) = 1\} = \lambda \Delta t + o(\Delta t), \quad \lambda > 0, \quad P\{\xi(t) \geq 2\} = o(\Delta t).$$

Ця властивість називається ординарністю, бо за час не більший Δt надходить не більше однієї вимоги (вимоги прибувають по одній).

Є системи, в яких вимоги прибувають по дві (наприклад органи реєстрації актів цивільного стану), але тут можна перейти до іншої вимоги, розглядаючи кожну пару вимог як одну вимогу. Будемо вважати, що на початку (при $t = t_1$) кількість вимог дорівнює нулю.

При такому підході зрозуміло, що інтервал часу між сусідніми вимогами розподілений за експоненціальним законом з параметром λ і число λ називається **інтенсивністю вимог**.

Будемо вважати, що час обслуговування вимоги кожним приладом розподілений експоненціально з параметром $\mu \geq 0$.

Якщо всі канали системи зайняті обслуговуванням вимог, то нова вимога може, в залежності від структури системи, стати в чергу на обслуговування або зникнути (загубитися).

Отже, **СМО можна класифікувати так:**

- одноканальні – багатоканальні;
- без втрат – з втратами вимог;
- з чергою – без черги (системи);
- з пріоритетним обслуговуванням – без пріоритетного обслуговування

Характеристиками системи є:

- абсолютна пропускна здатність системи (середня кількість);
- відносна пропускна здатність системи (відношення кількості вимог, які обслужили, до кількості всіх вимог);
- імовірність відмови в обслуговуванні;
- середня довжина черги;
- середній час простою каналу;
- середній час зайнятості каналу.

Відносно оптимізації виникають деякі конструктивні задачі: скільки потрібно приладів системи, щоб мінімізувати довжину черги або час перебування в черзі і т.д. Нас будуть цікавити не такі задачі.

Для характеристики системи, її опису може бути використаний такий підхід: під станом системи розумітимемо характеристику, яка відображає, у певний спосіб, кількість зайнятих каналів системи, кількість зайнятих місць у черзі (якщо така є).

Наприклад, якщо система складається з n пристроїв і черга відсутня, то під станами системи можемо розуміти таке:

0 – всі прилади вільні,

1 – один прилад зайнятий,

2 – два прилади зайняті,

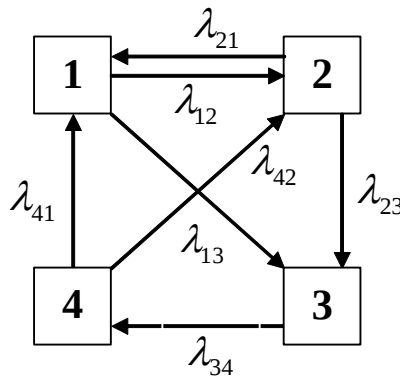
.....

n – n приладів зайняті.

З точки зору теорії ймовірності, цікаво простежити за тим, як система переходить з одного стану в інший.

Якби можна було знайти ймовірності переходу з будь-якого стану в будь-який інший стан і моменти часу таких переходів, то система була б описана повністю. Але це складна задача. Тому ставлять простішу задачу – знаходження ймовірностей $p_0(t)$, $p_1(t)$, $p_2(t), \dots, p_n(t), \dots$ того, що в момент часу t система перебувала в стані 0, 1 або n . Тоді можна було б знайти якісь характеристики системи. Поставимо завдання про знаходження таких ймовірностей: $p_0(t)$, $p_1(t)$, $p_2(t), \dots, p_n(t), \dots$ і т. д.

Нехай деяка система складається з чотирьох приладів, з'єднаних між собою в певний спосіб, λ_{ij} – інтенсивності переходу зі стану i в стан j .



Час знаходження системи в конкретному i -му стані (1, 2, 3, 4) перед переходом в j -тий стан розподілений експоненціально з параметром λ_{ij} .

Наприклад: знайдемо ймовірність того, що система перебуває в стані 4, тобто $p_4(t)$.

Знайдемо похідну від $p_4(t)$ по t .

Проаналізуємо, як система могла опинитися в стані 4: в момент часу t система знаходилась в стані 4 і за час Δt не перейшла ні в стан 1, ні в стан 2 або в момент часу t система знаходилась в стані 3 і за час Δt перейшла в стан 4.

Запишемо $p_4(t + \Delta t)$.

$$p_4(t + \Delta t) = p_4(t)(1 - \lambda_{41}\Delta t + o(\Delta t))(1 - \lambda_{42}\Delta t + o(\Delta t)) + \\ + p_3(t)(\lambda_{34}\Delta t + o(\Delta t)).$$

Звідси

$$p_4(t + \Delta t) - p_4(t) = -(\lambda_{41} + \lambda_{42})\Delta t \cdot p_4(t) + \lambda_{34}\Delta t \cdot p_3(t) + o(\Delta t).$$

Тут у доданок $o(\Delta t)$ об'єднані всі величини (доданки) вищого порядку малості відносно Δt . Тоді

$$\frac{dp_4(t)}{dt} = \lambda_{34}p_3(t) - (\lambda_{41} + \lambda_{42})p_4(t).$$

Аналогічно можна знайти

$$\frac{dp_1(t)}{dt} = -(\lambda_{12} + \lambda_{13})p_1(t) + \lambda_{21}p_2(t) + \lambda_{41}p_4(t),$$

$$\frac{dp_2(t)}{dt} = \lambda_{12}p_1(t) - (\lambda_{21} + \lambda_{23})p_2(t) + \lambda_{42}p_4(t),$$

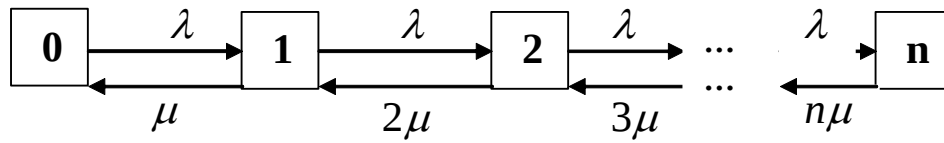
$$\frac{dp_3(t)}{dt} = \lambda_{13}p_1(t) + \lambda_{21}p_2(t) - \lambda_{34}p_3(t),$$

$$\frac{dp_4(t)}{dt} = \lambda_{34}p_3(t) - (\lambda_{41} + \lambda_{42})p_4(t).$$

Тут в останніх чотирьох рівностях біля вхідних стрілок знак “+”, а біля доданків з вихідними стрілками – знак “–”.

Розглянемо СМО з n приладами без черги. Всі прилади однакові. Інтенсивність вхідного потоку λ , інтенсивність обслуговування кожним приладом μ .

Стани: 0 – нуль приладів зайнято,
1 – один прилад зайнято і т. д.



Випишемо таку систему рівнянь:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t),$$

$$\frac{dp_1(t)}{dt} = \lambda p_0(t) - (\lambda + \mu) p_1(t) + 2\mu p_2(t),$$

$$\frac{dp_2(t)}{dt} = \lambda p_1(t) - (\lambda + 2\mu) p_2(t) + 3\mu p_3(t), \dots,$$

$$\frac{dp_{n-1}(t)}{dt} = \lambda p_{n-2}(t) - (\lambda + (n-1)\mu) p_{n-1}(t) + n\mu p_n(t),$$

$$\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - n\mu p_n(t),$$

$$\begin{cases} p_0(0) = 1, \\ p_1(0) = p_2(0) = \dots = p_n(0) = 0, \end{cases}$$

$$\sum_{i=0}^n p_i(t) = 1.$$

Ця система називається **системою рівнянь Ерланга***.

Можна довести, що існують границі $\lim_{t \rightarrow \infty} p_i(t) = p_i$, $i = \overline{0, n}$.

* Агнер Краруп Ерланг (Agner Krarup Erlang, 1878-1929) – датський математик, статистик та інженер.

Оскільки сума ймовірностей $\sum_{i=0}^n p_i(t) = 1$, то і $\sum_{i=0}^n p_i = 1$. Сталий

розвиток системи, який відповідає усталеному режиму роботи, знаходиться як розв'язок систем рівнянь

$$\begin{cases} -\lambda p_0 + \mu p_1 = 0, \\ \lambda p_0 - (\lambda + \mu) p_1 + 2\mu p_2 = 0, \\ \dots\dots\dots \\ \lambda p_{n-1} - n\mu p_n = 0, \\ p_0 + p_1 + p_2 + \dots + p_n = 1. \end{cases}$$

Розв'язуючи систему, з першого рівняння знаходимо p_1 через p_0 , з другого рівняння знаходимо p_2 через p_0 , і т.д.:

$$p_1 = \frac{\lambda}{\mu} p_0, \quad p_2 = \frac{\lambda^2}{2\mu^2} p_0, \quad p_3 = \frac{\lambda^3}{2 \cdot 3 \cdot \mu^3} p_0 \text{ і т.д.}$$

Позначимо $\frac{\lambda}{\mu} = \alpha$, тоді $p_i = \frac{\alpha^i}{i!} p_0$, ($i = \overline{1, n}$). Підставимо всі p_i ($i = \overline{1, n}$) в останнє рівняння системи:

$$p_0 = \left[1 + \frac{\alpha}{1} + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^n}{n!} \right]^{-1}.$$

Цими формулами визначається усталений режим роботи системи.

Знайдемо тепер імовірність відмови: $p_{\text{відмови}} = p_n$, бо всі прилади зайняті. Середня кількість зайнятих приладів дорівнює

$$\bar{n} = 0 \cdot p_0 + 1 \cdot p_1 + \dots + n \cdot p_n.$$

Всі інші характеристики, які можуть нас цікавити, в усталеному режимі роботи можуть бути також просто знайдені.

Розглянемо n -канальну систему з чергою. Є n каналів (приладів), усі прилади однотипні, інтенсивність виконання замовлень μ , у черзі передбачено m місць, вхідний потік вимог має інтенсивність λ . Система може перебувати у таких станах:

0 – всі прилади вільні, черга порожня;

1 – один прилад обслуговує вимогу, черга порожня;

2 – два прилади обслуговують заявку, черга порожня;

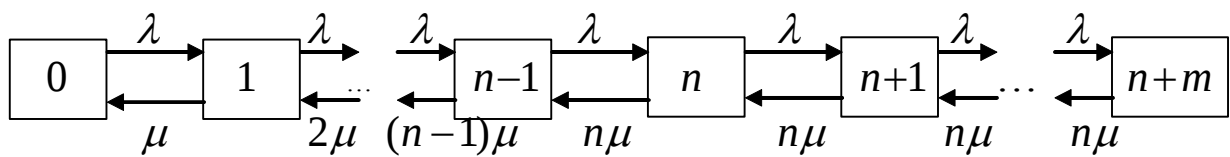
...

n – n приладів обслуговують заявку, черга порожня;

$n+1$ – n приладів обслуговують вимоги, одна вимога в черзі;

$n+2$ – n приладів обслуговують вимоги, дві вимоги перебувають в черзі;

$n+m$ – n приладів обслуговують вимоги, m вимог в черзі.



Треба знайти сталі ймовірності $p_0, p_1, \dots, p_n, p_{n+1}, \dots, p_{n+m}$, такі, що

$$\sum_{i=0}^{n+m} p_i = 1, \text{ де } p_i - \text{імовірність переходу в стан } i.$$

Випишемо систему лінійних рівнянь:

для стану 0: $-\lambda p_0 + \mu p_1 = 0,$

для стану 1: $\lambda p_0 - (\lambda + \mu) p_1 + 2\mu p_2 = 0,$

для стану 2: $\lambda p_1 - (\lambda + 2\mu) p_2 + 3\mu p_3 = 0,$

.....

для стану n : $\lambda p_{n-1} - (\lambda + n\mu) p_n + n\mu p_{n+1} = 0,$

для стану $n+1$: $\lambda p_n - (\lambda + n\mu)p_{n+1} + n\mu p_{n+2} = 0$,

.....

для стану $n+m$: $\lambda p_{n+m-1} - n\mu p_{n+m} = 0$,

$$\sum_{i=0}^{n+m} p_i = 1.$$

Позначивши $\frac{\lambda}{\mu} = \alpha$, маємо

$$p_1 = \frac{\alpha}{1!} p_0, \quad p_2 = \frac{\alpha^2}{2!} p_0, \dots, \quad p_n = \frac{\alpha^n}{n!} p_0,$$

$$p_{n+1} = \frac{\alpha^{n+1}}{n(n!)} p_0, \dots, \quad p_{n+m} = \frac{\alpha^{n+m}}{n^m(n!)} p_0.$$

Звідси, просумувавши, одержимо:

$$p_0 = \left[1 + \frac{\alpha}{1!} + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^n}{n!} + \frac{\alpha^{n+1}}{n(n!)} + \dots + \frac{\alpha^{n+m}}{n^m(n!)} \right]^{-1}.$$

Тоді задачі аналізу розв'язуються досить просто. Зокрема, такі:

1) імовірність відмови знаходиться так:

$$p_{\text{відмови}} = p_{n+m} \quad (\text{відмова у тому разі коли черга заповнена}),$$

2) відносна пропускна здатність: $1 - p_{\text{відмови}}$,

3) середня кількість вимог, які перебувають в черзі, в усталеному режимі – це випадкова величина з таким законом розподілу:

Можливі значення	0	1	2	...	m
Відповідні ймовірності	$\sum_{i=0}^n p_i$	p_{n+1}	p_{n+2}	...	p_{n+m}

Тоді середня кількість – це математичне сподівання і знаходиться просто.

4) середня кількість зайнятих каналів. Знову запишемо закон розподілу:

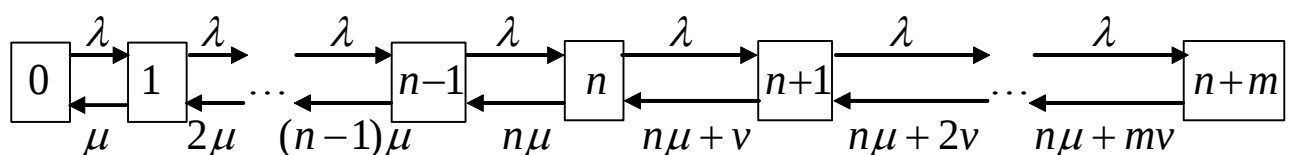
Можливі значення	0	1	2	...	n
Імовірності	p_0	p_1	p_2	...	$\sum_{i=0}^{n+m} p_i$

Середня кількість дорівнює математичному сподіванню цієї випадкової величини.

Якщо розглядати нескінченні черги, то замість сум треба розглядати ряди. Тоді ситуація ускладнюється, бо потрібно враховувати питання збіжності цих рядів.

Розглянемо систему з („нетерплячими заявками”) обмеженим часом перебування у черзі.

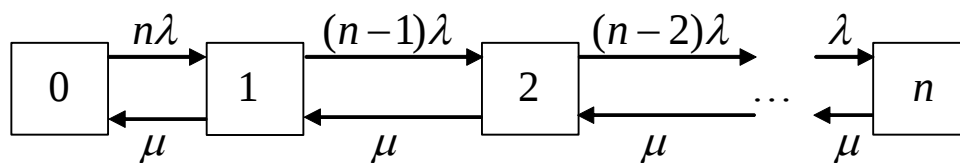
Є n каналів, m місць в черзі, але вимоги не можуть нескінченно довго перебувати в черзі. Нехай час перебування в черзі розподілений експоненціально з параметром ν (відповідна інтенсивність). Прикладом може слугувати обмеження в часі на перебування в черзі „до холодильника” продуктів харчування. У цьому прикладі сукупність станів така сама, як і в попередньому. Переходи здійснюються дещо інакше.



Розглянемо приклад замкнутої системи, для якої потік вимог внутрішній, тобто вимоги зароджуються в самій системі. Прикладом є налагодження верстатів на ткацькій фабриці. Всі прилади однотипні, виходять з ладу з інтенсивністю λ (час справної роботи розподілений експоненціально з параметром λ). Інтенсивність обслуговування μ , тобто час налагодження розподілений експоненціально з параметром μ .

Розглянемо такі стани:

- 0 – прилади працюють, наладник відпочиває;
- 1 – один прилад вийшов з ладу, наладник його обслуговує;
- 2 – два прилади вийшли з ладу, наладник обслуговує один прилад, один знаходиться у черзі;
- ...
- n – всі n приладів вийшли з ладу, наладник обслуговує один прилад, решта $n-1$ знаходяться в черзі.



Тут μ у нижньому рядку, бо наладник повертає в робочий стан лише один прилад.

ДОДАТКИ

Додаток А

Абсолютно неперервні функції

Припустимо, що на відрізку $[a, b]$ задана функція $f(x)$. Нехай інтервали $\sigma_i = (\alpha_i, \beta_i) \subset [a, b]$, $i = \overline{1, n}$.

Означення. Функція $f(x)$ називається *абсолютно неперервною* на $[a, b]$, якщо вона має таку властивість: для будь-якого $\varepsilon > 0$ існує $\delta > 0$, таке, що для довільної системи інтервалів $\sigma_1, \sigma_2, \dots, \sigma_n$, які попарно між собою не перетинаються і для яких $\sum_{j=1}^n |\sigma_j| < \delta$, виконується нерівність

$$\sum_{j=1}^n |f(\beta_j) - f(\alpha_j)| < \varepsilon.$$

Приклад 1. Абсолютно неперервною є функція, що задовольняє умову Ліпшиця.

Приклад 2. Абсолютно неперервною є функція

$$\Phi(x) = \int_a^x f(t) dt,$$

де $f \in L[a, b]$ – інтегровна за Лебегом на $[a, b]$.

Приклад 3. Побудована у §3 розділу 2 функція Кантора не є абсолютно неперервною.

Основні властивості абсолютно неперервних функцій:

- 1⁰. З абсолютної неперервності випливає рівномірна неперервність функції. Але обернений факт, взагалі кажучи, неправильний. Прикладом є функція Кантора – вона рівномірно неперервна, але не абсолютно неперервна.
- 2⁰. Якщо функція f – абсолютно неперервна, то функція αf – абсолютно неперервна.
- 3⁰. Якщо функції f і g – абсолютно неперервні, то функція $f + g$ – абсолютно неперервна.
- 4⁰. Всяка абсолютно неперервна функція є функцією обмеженої варіації.
- 5⁰. Для того, щоб на $[a, b]$ виконувалася рівність

$$\int_a^x F'(t)dt = F(x) - F(a),$$

необхідно і досить, щоб функція $F(x)$ була абсолютно неперервною.

Додаток Б

Інтеграл Стільтьєса

Нехай на інтервалі (a, b) визначені функція $f(x)$ і неспадна функція $F(x)$ з обмеженою варіацією. Будемо вважати, що функція $F(x)$ неперервна зліва. Якщо a і b скінченні, то розіб'ємо інтервал (a, b) точками $a = x_0 < x_1 < x_2 < \dots < x_n = b$ на скінченну кількість інтервалів (x_{i-1}, x_i) і утворимо суму

$$\sum_{i=1}^n f(\tilde{x}_i)[F(x_i) - F(x_{i-1})],$$

де $\tilde{x}_i$ — довільне число, вибране на сегменті (x_{i-1}, x_i) . Будемо збільшувати кількість точок поділу, одночасно наближаючи до нуля довжину максимального із сегментів. Якщо при цьому написана вище сума збігається до певної границі

$$J = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\tilde{x}_i)[F(x_i) - F(x_{i-1})],$$

то цю границю називають **інтегралом Стільтьєса** від функції $f(x)$ з інтегруючою функцією $F(x)$ і позначають символом

$$J = \int_a^b f(x) dF(x).$$

Невласний інтеграл Стільтьєса, коли проміжок інтегрування нескінченний, визначається так: розглядається інтеграл на довільному скінченному інтервалі (a, b) ; величини a і b довільним способом змушують прямувати до $-\infty$ і $+\infty$; якщо при цьому існує границя

$$\lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x) dF(x),$$

то ця границя називається інтегралом Стільтьєса від функції $f(x)$ по функції $F(x)$ на проміжку $(-\infty, +\infty)$ і позначається

$$\int_{-\infty}^{\infty} f(x) dF(x).$$

Всюди вважаємо, що інтеграл від функції $f(x)$ існує тоді і тільки тоді, коли існує інтеграл від $|f(x)|$ з тією ж інтегруючою функцією $F(x)$.

Якщо вважати (домовитися), що проміжок інтегрування має вигляд $[a, b]$, то можна показати, що

$$\int_a^b dF(x) = F(b) - F(a).$$

Зокрема, якщо $F(x)$ є функцією розподілу $F_{\xi}(x)$ випадкової величини ξ , то

$$\int_a^b dF_{\xi}(x) = F_{\xi}(b) - F_{\xi}(a) = P\{a \leq \xi < b\},$$

$$\int_{-\infty}^b dF_{\xi}(x) = F_{\xi}(b) = P\{\xi < b\}.$$

Якщо $F(x)$ має похідну $p(x)$ і є інтегралом від неї, то

$$\int_a^b f(x) dF(x) = \int_a^b f(x) p(x) dx$$

і інтеграл Стільтьєса зводиться до звичайного інтеграла Рімана.

Якщо зміна функції $F(x)$ відбувається тільки у точках $c_1, c_2, \dots, c_n, \dots$, то

$$\int_a^b f(x) dF(x) = \sum_{n=1}^{\infty} f(c_n) [F(c_n + 0) - F(c_n - 0)]$$

і інтеграл Стільтьєса перетворюється в ряд.

Перелічимо основні властивості інтеграла Стільтьєса:

1⁰. При $a = c_0 < c_1 < c_2 < \dots < c_n < c_{n+1} = b$

$$\int_a^b f(x) dF(x) = \sum_{i=0}^n \int_{c_i}^{c_{i+1}} f(x) dF(x).$$

2⁰. Постійний множник виноситься за знак інтеграла

$$\int_a^b c f(x) dF(x) = c \int_a^b f(x) dF(x).$$

3⁰. Інтеграл від суми функцій дорівнює сумі їх інтегралів

$$\int_a^b \sum_{i=1}^n f_i(x) dF(x) = \sum_{i=1}^n \int_a^b f_i(x) dF(x).$$

4⁰. Якщо $f(x) \geq 0$ і $b > a$, то

$$\int_a^b f(x) dF(x) \geq 0.$$

5⁰. Якщо $F_1(x)$ і $F_2(x)$ – монотонні функції з обмеженою змінною, а c_1 і c_2 – довільні константи, то

$$\int_a^b f(x) d[c_1 F_1(x) + c_2 F_2(x)] = c_1 \int_a^b f(x) dF_1(x) + c_2 \int_a^b f(x) dF_2(x).$$

6⁰. Якщо $F(x) = \int_c^x g(u) dG(u)$, де $g(u)$ – неперервна функція і

$G(u)$ – неспадна функція з обмеженою змінною, то

$$\int_a^b f(x) dF(x) = \int_a^b f(x) g(x) dG(x).$$

$$\text{Значення функції } \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3726	0,3712	0,3698
0,4	0,3683	0,3668	0,3652	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0395	0,0387	0,0379	0,0371	0,0363
2,2	0,0353	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001

Таблиця 2

Значення функції Лапласа $\Phi(x) \approx \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,0000	0,43	0,1664	0,86	0,3051	1,29	0,4015	1,72	0,4573	2,30	0,4893
0,01	0,0040	0,44	0,1700	0,87	0,3078	1,30	0,4032	1,73	0,4582	2,32	0,4898
0,02	0,0080	0,45	0,1736	0,88	0,3106	1,31	0,4049	1,74	0,4591	2,34	0,4904
0,03	0,0120	0,46	0,1772	0,89	0,3133	1,32	0,4066	1,75	0,4599	2,36	0,4909
0,04	0,0160	0,47	0,1808	0,90	0,3159	1,33	0,4082	1,76	0,4608	2,38	0,4913
0,05	0,0199	0,48	0,1844	0,91	0,3186	1,34	0,4099	1,77	0,4616	2,40	0,4918
0,06	0,0239	0,49	0,1879	0,92	0,3212	1,35	0,4115	1,78	0,4625	2,42	0,4922
0,07	0,0279	0,50	0,1915	0,93	0,3238	1,36	0,4131	1,79	0,4633	2,44	0,4927
0,08	0,0319	0,51	0,1950	0,94	0,3264	1,37	0,4147	1,80	0,4641	2,46	0,4931
0,09	0,0359	0,52	0,1985	0,95	0,3289	1,38	0,4162	1,81	0,4649	2,48	0,4934
0,10	0,0398	0,53	0,2019	0,96	0,3315	1,39	0,4177	1,82	0,4656	2,50	0,4938
0,11	0,0438	0,54	0,2054	0,97	0,3340	1,40	0,4192	1,83	0,4664	2,52	0,4941
0,12	0,0478	0,55	0,2088	0,98	0,3365	1,41	0,4207	1,84	0,4671	2,54	0,4945
0,13	0,0517	0,56	0,2123	0,99	0,3389	1,42	0,4222	1,85	0,4678	2,56	0,4948
0,14	0,0557	0,57	0,2157	1,00	0,3413	1,43	0,4236	1,86	0,4686	2,58	0,4951
0,15	0,0596	0,58	0,2190	1,01	0,3438	1,44	0,4251	1,87	0,4693	2,60	0,4953
0,16	0,0636	0,59	0,2224	1,02	0,3461	1,45	0,4265	1,88	0,4699	2,62	0,4956
0,17	0,0675	0,60	0,2257	1,03	0,3485	1,46	0,4279	1,89	0,4706	2,64	0,4959
0,18	0,0714	0,61	0,2291	1,04	0,3508	1,47	0,4292	1,90	0,4713	2,66	0,4961
0,19	0,0753	0,62	0,2324	1,05	0,3531	1,48	0,4306	1,91	0,4719	2,68	0,4963
0,20	0,0793	0,63	0,2357	1,06	0,3554	1,49	0,4319	1,92	0,4726	2,70	0,4965
0,21	0,0832	0,64	0,2389	1,07	0,3577	1,50	0,4332	1,93	0,4732	2,72	0,4967
0,22	0,0871	0,65	0,2422	1,08	0,3599	1,51	0,4345	1,94	0,4738	2,74	0,4969
0,23	0,0910	0,66	0,2454	1,09	0,3621	1,52	0,4357	1,95	0,4744	2,76	0,4971
0,24	0,0948	0,67	0,2486	1,10	0,3643	1,53	0,4370	1,96	0,4750	2,78	0,4973
0,25	0,0987	0,68	0,2517	1,11	0,3665	1,54	0,4382	1,97	0,4756	2,80	0,4974
0,26	0,1026	0,69	0,2549	1,12	0,3686	1,55	0,4394	1,98	0,4761	2,82	0,4976
0,27	0,1064	0,70	0,2580	1,13	0,3708	1,56	0,4406	1,99	0,4767	2,84	0,4977
0,28	0,1103	0,71	0,2611	1,14	0,3729	1,57	0,4418	2,00	0,4772	2,86	0,4979
0,29	0,1141	0,72	0,2642	1,15	0,3749	1,58	0,4429	2,02	0,4783	2,88	0,4980
0,30	0,1179	0,73	0,2673	1,16	0,3770	1,59	0,4441	2,04	0,4793	2,90	0,4981
0,31	0,1217	0,74	0,2703	1,17	0,3790	1,60	0,4452	2,06	0,4803	2,92	0,4982
0,32	0,1255	0,75	0,2734	1,18	0,3810	1,61	0,4463	2,08	0,4812	2,94	0,4984
0,33	0,1293	0,76	0,2764	1,19	0,3830	1,62	0,4474	2,10	0,4821	2,96	0,4985
0,34	0,1331	0,77	0,2794	1,20	0,3849	1,63	0,4484	2,12	0,4830	2,98	0,4986
0,35	0,1368	0,78	0,2823	1,21	0,3869	1,64	0,4495	2,14	0,4838	3,00	0,49865
0,36	0,1406	0,79	0,2852	1,22	0,3883	1,65	0,4505	2,16	0,4846	3,20	0,49931
0,37	0,1443	0,80	0,2881	1,23	0,3907	1,66	0,4515	2,18	0,4854	3,40	0,49966
0,38	0,1480	0,81	0,2910	1,24	0,3925	1,67	0,4525	2,20	0,4861	3,60	0,499841
0,39	0,1517	0,82	0,2939	1,25	0,3944	1,68	0,4535	2,22	0,4868	3,80	0,499928
0,40	0,1554	0,83	0,2967	1,26	0,3962	1,69	0,4545	2,24	0,4875	4,00	0,499968
0,41	0,1591	0,84	0,2995	1,27	0,3980	1,70	0,4554	2,26	0,4881	4,50	0,499997
0,42	0,1628	0,85	0,3023	1,28	0,3997	1,71	0,4564	2,28	0,4887	5,00	0,499997

Таблиця 3

Імовірність $P_{\lambda}(\xi = m) = \frac{\lambda^m}{m!} e^{-\lambda}$ розподілу Пуассона

$\lambda \backslash m$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	0,90484	0,81873	0,74082	0,67032	0,60653	0,54881	0,49659	0,44933	0,40657
1	0,09048	0,16375	0,22225	0,26813	0,30327	0,32929	0,34761	0,35946	0,36591
2	0,00452	0,01637	0,03334	0,05363	0,07582	0,09879	0,12166	0,14379	0,16466
3	0,00015	0,00109	0,00333	0,00715	0,01264	0,01976	0,02839	0,03834	0,04940
4		0,00005	0,00025	0,00072	0,00158	0,00296	0,00497	0,00767	0,01111
5			0,00002	0,00006	0,00016	0,00036	0,00070	0,00123	0,00200
6					0,00001	0,00004	0,00008	0,00016	0,00030
7								0,00002	0,00004

$\lambda \backslash m$	1	2	3	4	5	6	7	8	9	10
0	0,36788	0,13534	0,04979	0,01832	0,00674	0,00248	0,00091	0,00034	0,00012	0,00005
1	0,36788	0,27067	0,14936	0,07326	0,03369	0,01487	0,00638	0,00268	0,00111	0,00045
2	0,18394	0,27067	0,22404	0,14653	0,08422	0,04462	0,02234	0,01073	0,00500	0,00227
3	0,06131	0,18045	0,22404	0,19537	0,14037	0,08924	0,05213	0,02863	0,01499	0,00757
4	0,01533	0,09022	0,16803	0,19537	0,17547	0,13385	0,09123	0,05725	0,03374	0,01892
5	0,00307	0,03609	0,10082	0,15629	0,17547	0,16062	0,12772	0,09160	0,06073	0,03783
6	0,00051	0,01203	0,05041	0,10420	0,14622	0,16062	0,14900	0,12214	0,09109	0,06306
7	0,00007	0,00344	0,02160	0,05954	0,10444	0,13768	0,14900	0,13959	0,11712	0,09008
8		0,00086	0,00810	0,2977	0,06528	0,10326	0,13038	0,13959	0,13176	0,11260
9		0,00019	0,00270	0,01323	0,03627	0,06884	0,10140	0,12408	0,13176	0,12511
10		0,00004	0,00081	0,00529	0,01813	0,04130	0,07098	0,09926	0,11858	0,12511
11			0,00022	0,00192	0,00824	0,02253	0,04517	0,07219	0,09702	0,11374
12			0,00006	0,00064	0,00343	0,01126	0,02635	0,04813	0,07277	0,09478
13				0,00020	0,00132	0,00520	0,01419	0,02962	0,05038	0,07291
14				0,00006	0,00047	0,00223	0,00709	0,01692	0,03238	0,05208
15				0,00002	0,00016	0,00089	0,00331	0,00903	0,01943	0,3472
16					0,00005	0,00033	0,00145	0,00451	0,01093	0,2170
17					0,00001	0,00012	0,00060	0,00212	0,00579	0,01276
18						0,00004	0,00023	0,00094	0,00289	0,00709
19						0,00001	0,00009	0,00040	0,00137	0,00373
20							0,00003	0,00016	0,00062	0,00187
21								0,00006	0,00026	0,00089
22								0,00002	0,00011	0,00040
23									0,00004	0,00018
24									0,00002	0,00007
25										0,00003
26										0,00001

Продовження табл. 3

$\lambda \backslash m$	11	12	13	14	15	16	17	18	19	20
0	0,00002									
1	0,00018	0,00007	0,00003	0,00001						
2	0,00101	0,00044	0,00019	0,00008	0,00003	0,00001				
3	0,00370	0,00177	0,00083	0,00038	0,00017	0,00008	0,00003	0,00001		
4	0,01019	0,00531	0,00269	0,00133	0,00065	0,00031	0,00014	0,00007	0,00003	0,00001
5	0,02242	0,01274	0,00699	0,00373	0,00194	0,00098	0,00049	0,00024	0,00012	0,00005
6	0,04109	0,02548	0,01515	0,00870	0,00484	0,00262	0,00139	0,00072	0,00037	0,00018
7	0,06458	0,04368	0,02814	0,01739	0,01037	0,00599	0,00337	0,00185	0,00099	0,00052
8	0,08879	0,06552	0,04573	0,03044	0,01944	0,01199	0,00716	0,00416	0,00236	0,00131
9	0,10853	0,08736	0,06605	0,04734	0,03241	0,02131	0,01353	0,00833	0,00498	0,00291
10	0,11938	0,10484	0,08587	0,06628	0,04861	0,03410	0,02300	0,01499	0,00947	0,00582
11	0,11938	0,11437	0,10148	0,08436	0,06629	0,04960	0,03554	0,02452	0,01635	0,01058
12	0,10943	0,11437	0,10994	0,09842	0,08286	0,06613	0,05036	0,03678	0,02589	0,01763
13	0,09259	0,10557	0,10994	0,10599	0,09561	0,08139	0,06585	0,05093	0,03784	0,02712
14	0,07275	0,09049	0,10209	0,10599	0,10244	0,09302	0,07996	0,06548	0,05135	0,03874
15	0,05335	0,07239	0,08848	0,09892	0,10244	0,09922	0,09062	0,07858	0,06504	0,05165
16	0,03668	0,05429	0,07189	0,08656	0,09603	0,09922	0,09628	0,08840	0,07724	0,06456
17	0,02373	0,03832	0,05497	0,07128	0,08474	0,09338	0,09628	0,09360	0,08633	0,07595
18	0,01450	0,02555	0,03970	0,05544	0,07061	0,08301	0,09094	0,09360	0,09112	0,08439
19	0,00840	0,01614	0,02716	0,04085	0,05575	0,06990	0,08136	0,08867	0,09112	0,08884
20	0,00462	0,00968	0,01766	0,02860	0,04181	0,05592	0,06916	0,07980	0,08657	0,08884
21	0,00242	0,00553	0,01093	0,01906	0,02986	0,04261	0,05599	0,06840	0,07832	0,08461
22	0,00121	0,00302	0,00646	0,01213	0,02036	0,03099	0,04326	0,05597	0,06774	0,07691
23	0,00058	0,00157	0,00365	0,00738	0,01328	0,02156	0,03198	0,04380	0,05588	0,06688
24	0,00027	0,00079	0,00198	0,00431	0,00830	0,01437	0,02265	0,03285	0,04424	0,05573
25	0,00012	0,00038	0,00103	0,00241	0,00498	0,00920	0,01540	0,02365	0,03362	0,04459
26	0,00005	0,00017	0,00051	0,00130	0,00287	0,00566	0,01007	0,01637	0,02457	0,03430
27	0,00002	0,00008	0,00025	0,00067	0,00160	0,00335	0,00634	0,01092	0,01729	0,02541
28		0,00003	0,00011	0,00034	0,00086	0,00192	0,00385	0,00702	0,01173	0,01815
29		0,00001	0,00005	0,00016	0,00044	0,00106	0,00226	0,00436	0,00769	0,01252
30			0,00002	0,00008	0,00022	0,00056	0,00128	0,00261	0,00487	0,00834
31				0,00003	0,00011	0,00029	0,00070	0,00152	0,00298	0,00538
32				0,00001	0,00005	0,00015	0,00037	0,00085	0,00177	0,00336
33					0,00002	0,00007	0,00019	0,00047	0,00102	0,00204
34					0,00001	0,00003	0,00010	0,00025	0,00057	0,00120
35						0,00002	0,00005	0,00013	0,00031	0,00069
36							0,00002	0,00006	0,00016	0,00038
37							0,00001	0,00003	0,00008	0,00021
38								0,00001	0,00004	0,00011
39									0,00002	0,00006
40										0,00003

Таблиця 4

Значення $t_\gamma = t(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,00	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,001	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Таблиця 5

Значення $q = q(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Таблиця 6

Значення ймовірності $P\{U \geq \chi^2\}$ χ^2 -розподілу Пірсона

$\chi^2 \backslash r$	1	2	3	4	5	6	7	8	9	10
1	0.317311	0.606531	0.801252	0.909796	0.962566	0.985612	0.994829	0.998248	0.999438	0.999828
2	0.157299	0.367879	0.572407	0.735759	0.849145	0.919699	0.959840	0.981012	0.991468	0.996340
3	0.083265	0.223130	0.391625	0.557825	0.699986	0.808847	0.885002	0.934358	0.964295	0.981424
4	0.045500	0.135335	0.261464	0.406006	0.549416	0.676676	0.779777	0.857123	0.911413	0.947347
5	0.025347	0.082085	0.171797	0.287297	0.415880	0.543813	0.659963	0.757576	0.834308	0.891178
6	0.014306	0.049787	0.111610	0.199148	0.306219	0.423190	0.539749	0.647232	0.739918	0.815263
7	0.008151	0.030197	0.071898	0.135888	0.220640	0.320847	0.428880	0.536633	0.637119	0.725445
8	0.004678	0.018316	0.046012	0.091578	0.156236	0.238103	0.332594	0.433470	0.534146	0.628837
9	0.002700	0.011109	0.029291	0.061099	0.109064	0.173578	0.252656	0.342296	0.437274	0.532104
10	0.001565	0.006738	0.018566	0.040428	0.075235	0.124652	0.188573	0.265026	0.350485	0.440493
11	0.000911	0.004087	0.011726	0.026564	0.051380	0.088376	0.138619	0.201699	0.275709	0.357518
12	0.000532	0.002479	0.007383	0.017351	0.034788	0.061969	0.100559	0.151204	0.213309	0.285057
13	0.000311	0.001503	0.004637	0.011276	0.023379	0.043036	0.072108	0.111850	0.162606	0.223672
14	0.000183	0.000912	0.002905	0.007295	0.015609	0.029636	0.051181	0.081765	0.122325	0.172992
15	0.000108	0.000553	0.001817	0.004701	0.010362	0.020257	0.035999	0.059145	0.090936	0.132062
16	0.000063	0.000335	0.001134	0.003019	0.006844	0.013754	0.025116	0.042380	0.066882	0.099632
17	0.000037	0.000203	0.000707	0.001933	0.004500	0.009283	0.017396	0.030109	0.048716	0.074364
18	0.000022	0.000123	0.000440	0.001234	0.002946	0.006232	0.011970	0.021226	0.035174	0.054964
19	0.000013	0.000075	0.000273	0.000786	0.001922	0.004164	0.008187	0.014860	0.025193	0.040263
20	0.000008	0.000045	0.000170	0.000499	0.001250	0.002769	0.005570	0.010336	0.017912	0.029253
21	0.000005	0.000028	0.000105	0.000317	0.000810	0.001835	0.003770	0.007147	0.012650	0.021094
22	0.000003	0.000017	0.000065	0.000200	0.000524	0.001211	0.002540	0.004916	0.008879	0.015105
23	0.000002	0.000010	0.000040	0.000127	0.000338	0.000796	0.001705	0.003364	0.006196	0.010747
24	0.000001	0.000006	0.000025	0.000080	0.000217	0.000522	0.001139	0.002292	0.004301	0.007600
25	0.000001	0.000004	0.000015	0.000050	0.000139	0.000341	0.000759	0.001555	0.002971	0.005346
26	0.000000	0.000002	0.000010	0.000032	0.000089	0.000223	0.000504	0.001050	0.002043	0.003740
27	0.000000	0.000001	0.000006	0.000020	0.000057	0.000145	0.000333	0.000707	0.001399	0.002604
28	0.000000	0.000001	0.000004	0.000012	0.000036	0.000094	0.000220	0.000474	0.000954	0.001805
29	0.000000	0.000001	0.000002	0.000008	0.000023	0.000061	0.000145	0.000317	0.000648	0.001246
30	0.000000	0.000000	0.000001	0.000005	0.000015	0.000039	0.000095	0.000211	0.000439	0.000857
31	0.000000	0.000000	0.000001	0.000003	0.000009	0.000025	0.000062	0.000141	0.000296	0.000587
32	0.000000	0.000000	0.000001	0.000002	0.000006	0.000016	0.000041	0.000093	0.000199	0.000400
33	0.000000	0.000000	0.000000	0.000001	0.000004	0.000010	0.000026	0.000062	0.000134	0.000272
34	0.000000	0.000000	0.000000	0.000001	0.000002	0.000007	0.000017	0.000041	0.000089	0.000185
35	0.000000	0.000000	0.000000	0.000000	0.000002	0.000004	0.000011	0.000027	0.000060	0.000125
36	0.000000	0.000000	0.000000	0.000000	0.000001	0.000003	0.000007	0.000018	0.000040	0.000084
37	0.000000	0.000000	0.000000	0.000000	0.000001	0.000002	0.000005	0.000012	0.000026	0.000057
38	0.000000	0.000000	0.000000	0.000000	0.000000	0.000001	0.000003	0.000008	0.000017	0.000038
39	0.000000	0.000000	0.000000	0.000000	0.000000	0.000001	0.000002	0.000005	0.000012	0.000025
40	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000001	0.000003	0.000008	0.000017

Таблиця 7

$$\text{Значення функції } P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda^2}$$

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
≤0,29	1	0,63	0,8222	0,97	0,3036	1,31	0,0646	1,65	0,0086
0,3	1	0,64	0,8073	0,98	0,2921	1,32	0,0613	1,66	0,0081
0,31	1	0,65	0,792	0,99	0,2809	1,33	0,0582	1,67	0,0076
0,32	1	0,66	0,7764	1	0,27	1,34	0,0551	1,68	0,0071
0,33	0,9999	0,67	0,7604	1,01	0,2594	1,35	0,0522	1,69	0,0066
0,34	0,9998	0,68	0,7442	1,02	0,2492	1,36	0,0495	1,7	0,0062
0,35	0,9997	0,69	0,7278	1,03	0,2392	1,37	0,0469	1,71	0,0058
0,36	0,9995	0,7	0,7112	1,04	0,2296	1,38	0,0444	1,72	0,0054
0,37	0,9992	0,71	0,6945	1,05	0,2202	1,39	0,042	1,73	0,005
0,38	0,9987	0,72	0,6777	1,06	0,2111	1,4	0,0397	1,74	0,0047
0,39	0,9981	0,73	0,6609	1,07	0,2024	1,41	0,0375	1,75	0,0044
0,4	0,9972	0,74	0,644	1,08	0,1939	1,42	0,0354	1,76	0,0041
0,41	0,996	0,75	0,6272	1,09	0,1857	1,43	0,0335	1,77	0,0038
0,42	0,9945	0,76	0,6104	1,1	0,1777	1,44	0,0316	1,78	0,0035
0,43	0,9926	0,77	0,5936	1,11	0,17	1,45	0,0298	1,79	0,0033
0,44	0,9903	0,78	0,577	1,12	0,1626	1,46	0,0282	1,8	0,0031
0,45	0,9874	0,79	0,5605	1,13	0,1555	1,47	0,0266	1,81	0,0029
0,46	0,984	0,8	0,5441	1,14	0,1486	1,48	0,025	1,82	0,0027
0,47	0,98	0,81	0,528	1,15	0,142	1,49	0,0236	1,83	0,0025
0,48	0,9753	0,82	0,512	1,16	0,1356	1,5	0,0222	1,84	0,0023
0,49	0,97	0,83	0,4962	1,17	0,1294	1,51	0,0209	1,85	0,0021
0,5	0,9639	0,84	0,4806	1,18	0,1235	1,52	0,0197	1,86	0,002
0,51	0,9572	0,85	0,4653	1,19	0,1177	1,53	0,0185	1,87	0,0019
0,52	0,9497	0,86	0,4503	1,2	0,1122	1,54	0,0174	1,88	0,0017
0,53	0,9415	0,87	0,4355	1,21	0,107	1,55	0,0164	1,89	0,0016
0,54	0,9325	0,88	0,4209	1,22	0,1019	1,56	0,0154	1,9	0,0015
0,55	0,9228	0,89	0,4067	1,23	0,097	1,57	0,0145	1,91	0,0014
0,56	0,9124	0,9	0,3927	1,24	0,0924	1,58	0,0136	1,92	0,0013
0,57	0,9013	0,91	0,3791	1,25	0,0879	1,59	0,0127	1,93	0,0012
0,58	0,8896	0,92	0,3657	1,26	0,0836	1,6	0,012	1,94	0,0011
0,59	0,8772	0,93	0,3527	1,27	0,0794	1,61	0,0112	1,95	0,001
0,6	0,8643	0,94	0,3399	1,28	0,0755	1,62	0,0105	1,96	0,0009
0,61	0,8508	0,95	0,3275	1,29	0,0717	1,63	0,0098	1,97	0,0009
0,62	0,8368	0,96	0,3154	1,3	0,0681	1,64	0,0092	1,98	0,0008

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
1,99	0,0007	2,2	0,0001	2,41	0,000018
2	0,0007	2,21	0,0001	2,42	0,000016
2,01	0,0006	2,22	0,0001	2,43	0,000014
2,02	0,0006	2,23	0,0001	2,44	0,000013
2,03	0,0005	2,24	0,0001	2,45	0,000012
2,04	0,0005	2,25	0,0001	2,46	0,000011
2,05	0,0004	2,26	0,0001	2,47	0,00001
2,06	0,0004	2,27	0,0001	2,48	0,000009
2,07	0,0004	2,28	0,0001	2,49	0,000008
2,08	0,0004	2,29	0,0001	2,5	0,0000075
2,09	0,0003	2,3	0,0001	2,55	0,0000044
2,1	0,0003	2,31	0,000046	2,6	0,0000026
2,11	0,0003	2,32	0,000042	2,65	0,0000016
2,12	0,0002	2,33	0,000038	2,7	0,000001
2,13	0,0002	2,34	0,000035	2,75	0,0000006
2,14	0,0002	2,35	0,000032	2,8	0,0000003
2,15	0,0002	2,36	0,00003	2,85	0,00000018
2,16	0,0002	2,37	0,000027	2,9	0,0000001
2,17	0,0002	2,38	0,000024	2,95	0,00000006
2,18	0,0001	2,39	0,000022	3	0,00000003
2,19	0,0001	2,4	0,00002		

СПИСОК ЛІТЕРАТУРИ

1. Агапова І. С., Бондаренко М. Ф., Дікареєв В. А., Семенець В. В. Збірник задач з теорії ймовірностей з розв'язками : навч. посібник / за ред. М. Ф. Бондаренка. Харків : ХНУРЕ, 2010. 356 с.
2. Барковський В. В. Теорія ймовірностей та математична статистика. / В. В. Барковський, Н. В. Барковська, О. К. Лопатін. Київ : Центр учбової літератури, 2010. 424 с.
3. Васильків І. М. Основи теорії ймовірностей і математичної статистики : навч. посібник. Львів : ЛНУ імені Івана Франка, 2020. 184 с.
4. Гнеденко Б. В. Курс теорії ймовірностей : навч. посібник / Б. В. Гнеденко Київ : Видавничо-поліграфічний центр „Київський університет”, 2010. 464 с.
5. Голомозий В.В. Збірник задач з теорії ймовірностей та математичної статистики : навч. посібник / В. В. Голомозий, М. В. Карташов, К. В. Ральченко. Київ : Видавничо-поліграфічний центр „Київський університет”, 2015. 366 с.
6. Жильцов О. Б. Теорія ймовірностей та математична статистика у прикладах і задачах : навч. посіб. для студентів немат. спец. ВНЗ / О. Б. Жильцов. Київ : Київ. ун–т ім. Бориса Грінченка. 2015. 335 с.
7. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посібник. У 2 ч. Ч.1. Теорія ймовірностей. / В. І. Жлуктенко, С. І. Наконечний. Вид. 2-ге, без змін. Київ : КНЕУ, 2007. 304 с.
8. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посібник. У 2 ч. Ч.2. Математична статистика. /

- В. І. Жлуктенко, С. І. Наконечний. Вид. 2-ге, без змін. Київ : КНЕУ, 2007. 336 с.
9. Кармелюк Г. І. Теорія ймовірностей та математична статистика. Посібник з розв'язування задач : навч. посібник / Г. І. Кармелюк. Київ : Центр учбової літератури, 2007. 576 с.
 10. Карташов М. В. Теорія ймовірностей і математична статистика : підручник. Київ : Видавничо-поліграфічний центр „Київський університет”, 2009. 479 с.
 11. Костробій П. П. Теорія ймовірностей. Львів : Видавництво „Растр-7”, 2016. 260 с.
 12. Сеньо П. С. Теорія ймовірностей та математична статистика : підручник. / П. С. Сеньо. 2-ге вид. Київ : Знання, 2007. 556 с.
 13. Слюсарчук Ю. М. Теорія ймовірностей, математична статистика та імовірнісні процеси : навч. посібник / Ю. М. Слюсарчук, Й. Я. Хром'як, Л. Л. Джавала, В. М. Цимбал. Львів : Вид-во Львів. політехніки, 2015. 364 с.
 14. Цибенко О. С., Крищук М. Г., Тарасевич Ю. Я. Збірник задач з теорії ймовірностей : навч. посібник. Київ : Національний тех.-нічний університет України „Київський політехнічний інститут”, 2016. 210 с.
 15. Черняк О.І. Теорія ймовірностей та математична статистика. Практикум : навчальний посібник для студентів закладів вищої освіти. / О. І. Черняк, Т. В. Кравець, О. І. Ляшенко, Л. М. Буяк, О. С. Башуцька. Тернопіль : ТНЕУ, 2019. 251 с.

ЗМІСТ

ВСТУП	3
ВІД АВТОРА	4
РОЗДІЛ 1. ВИПАДКОВІ ПОДІЇ	5
§1. Стохастичний експеримент. Випадкова подія. Частота випадкової події	5
§2. Класичне означення ймовірності	12
§3. Геометрична ймовірність	16
§4. Класи подій	18
§5. Аксиоматика теорії ймовірності	23
§6. Властивості ймовірності	25
§7. Умовні ймовірності	31
§8. Формули повної ймовірності, Бейєса	39
§9. Схема незалежних випробувань. Закон Бернуллі (біноміальний закон)	43
п. 9.1. Найбільш імовірне число	44
п. 9.2. Біноміальний і геометричний розподіли	47
§10. Граничні теореми у схемі незалежних випробувань	50
п. 10.1 Локальна теорема Муавра – Лапласа	50
п. 10.2. Інтегральна теорема Муавра – Лапласа (наслідок локальної теореми)	55
п. 10.3 Теорема Пуассона (теорема про рідкісні події)	58

РОЗДІЛ 2. ВИПАДКОВІ ВЕЛИЧИНИ	60
§1. Поняття випадкової величини	60
§2. Властивості функції розподілу випадкової величини	63
§3. Класифікація випадкових величин	77
п. 3.1. Властивості щільності розподілу	81
п. 3.2. Показниковий розподіл	84
п. 3.3. Розподіл Коші	87
§4. Векторні випадкові величини	89
§5. Щільність сумісного розподілу. Розподіл суми випадкових величин	97

РОЗДІЛ 3. ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВИХ ВЕЛИЧИН	109
§1. Числові характеристики випадкових величин	110
§2. Властивості числових характеристик	120
п. 2.1. Властивості математичного сподівання	120
п. 2.2. Властивості дисперсії	127
п. 2.3. Характеристики положення і форми розподілу випадкової величини	133
п. 2.4. Кореляція двох випадкових величин	138
§3. Характеристична функція випадкової величини	145
§4. Послідовності випадкових величин	154
§5 Закон великих чисел	160

§6 Центральна гранична теорема	168
РОЗДІЛ 4. МАТЕМАТИЧНА СТАТИСТИКА	172
§1. Предмет і задачі математичної статистики	172
§2. Первинна обробка статистичної інформації	182
§3. Оцінювання числових характеристик	185
§4. Довірчий інтервал	188
§5. Критерії згоди	202
§6. Принцип найбільшої правдоподібності. Метод найменших квадратів згладжування результатів експерименту	208
§7. Згладжування результатів спостережень за двовимірною випадковою величиною. Регресія. Лінійна регресія	213
РОЗДІЛ 5. ЕЛЕМЕНТИ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ (ФУНКЦІЙ)	219
§1. Основні поняття	219
§2. Процес Пуассона	224
§3. Процес Вінера (вінерівський процес)	228
§4. Стохастичний інтеграл Іто. Стохастичне диференціальне рівняння	231
§5. Канонічне представлення стаціонарного випадкового процесу	238
§6. Поняття про марковський процес (процес Маркова)	248
§7. Аналіз випадкових процесів у системах масового обслуговування	253

ДОДАТКИ	264
Додаток А. Абсолютно неперервні функції	264
Додаток Б. Інтеграл Стільтьєса	266
Додаток В. Таблиці	269
СПИСОК ЛІТЕРАТУРИ	278
ЗМІСТ	280

Навчальне видання

Василь Йосипович **Кушнірчук**

ТЕОРІЯ ЙМОВІРНОСТЕЙ

Підручник

Літературне редагування ***О.В. Колодій***

Технічне редагування та дизайн обкладинки ***Г.К. Чорасва***

Підписано до друку 22.12.2025. Формат 60х84/16.

Папір офсетний. Друк цифровий. Умов.-друк. арк. 15,5.

Обл.-вид. арк. 16,7. Зам. Н-091.

Видавництво та друкарня Чернівецького національного університету.

58002, Чернівці, вул. Коцюбинського, 2.

e-mail: ruta@chnu.edu.ua

Свідоцтво суб'єкта видавничої справи ДК № 891 від 08.04.2002.