

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

СТОХАСТИЧНІ ОСНОВИ АНАЛІЗУ ДАНИХ

Конспект лекцій

Навчальний посібник

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за освітньою програмою
«Промисловий маркетинг»
спеціальності D5 Маркетинг*

Укладачі: С. О. Солнцев, О. В. Черненко, О. Є. Кофанов

Електронне мережеве навчальне видання

Київ

КПІ ім. Ігоря Сікорського

2025

УДК 519.2:330(075.8)

T11

Укладачі: *Солнцев Сергій Олексійович*
доктор фізико-математичних наук, професор,
професор кафедри промислового маркетингу
Черненко Оксана Володимирівна
кандидат економічних наук, доцент
Кофанов Олексій Євгенович
кандидат економічних наук, кандидат технічних наук, доцент,
доцент кафедри промислового маркетингу

Рецензент: *Капустян В. О.*, доктор фізико-математичних наук,
професор, професор кафедри економічної кібернетики
КПІ ім. Ігоря Сікорського

Відповідальний редактор *Жигалкевич Ж. М.*, доктор економічних наук, професор,
професор кафедри промислового маркетингу

*Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 2 від 06.11.2025 р.)
за поданням Вченої ради факультету менеджменту та маркетингу
(протокол № 2 від 29.09.2025 р.).*

T11 Стохастичні основи аналізу даних. Конспект лекцій [Електронний ресурс] :
навч. посіб. для здобувачів ступеня бакалавра за освіт. програмою
«Промисловий маркетинг» спеціальності D5 Маркетинг / КПІ ім. Ігоря
Сікорського ; уклад.: С. О. Солнцев, О. В. Черненко,
О. Є. Кофанов. – Електрон. текст. дані (1 файл). – Київ : КПІ ім. Ігоря
Сікорського, 2025. – 160 с.

Посібник містить теоретичні положення та методи стохастичного аналізу, необхідні для моделювання процесів і роботи з даними в умовах невизначеності. Розглянуто базові ймовірнісні моделі, застосування комбінаторики, байєсівські методи оновлення прогнозів, стохастичні моделі дискретних і неперервних ознак, числові характеристики розподілів, а також методи узагальнення вибірових даних, статистичного оцінювання та перевірки гіпотез. Матеріал подано з акцентом на розв'язання прикладних аналітичних задач і прийняття рішень на основі даних.

УДК 519.2:330(075.8)

Реєстр. № 25/26-044. Обсяг 8,0 авт. арк.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
проспект Берестейський, 37, м. Київ, 03056
<https://kpi.ua>

Свідоцтво про внесення до Державного реєстру видавців, виготовлювачів
і розповсюджувачів видавничої продукції ДК № 5354 від 25.05.2017 р.

© С. О. Солнцев, О. В. Черненко, О. Є. Кофанов, 2025
© КПІ ім. Ігоря Сікорського, 2025

ЗМІСТ

ВСТУП	5
РОЗДІЛ 1. ВИПАДКОВІ ПОДІЇ ТА ЇХ МОДЕЛЮВАННЯ.....	7
Тема 1. Випадкові події та міри невизначеності в базових імовірнісних моделях	7
Лекційне заняття № 1	7
Тема 2. Застосування комбінаторики для обчислення ймовірностей	24
Лекційне заняття № 2	24
Тема 3. Умовна ймовірність та незалежність подій	26
Лекційне заняття № 3	26
Тема 4. Формула повної ймовірності у моделюванні складних подій.....	29
Лекційне заняття № 4	29
Тема 5. Формула Байєса та переоцінка ймовірностей на основі нових даних.....	32
Лекційне заняття № 5	32
Тема 6. Моделювання послідовних випробувань за схемою Бернуллі... 36	
Лекційне заняття № 6	36
РОЗДІЛ 2. ЙМОВІРНІСНІ МОДЕЛІ ДАНИХ (ВИПАДКОВІ ВЕЛИЧИНИ).....	39
Тема 7. Стохастичні моделі для дискретних даних	39
Лекційне заняття № 7	39
Тема 8. Стохастичні моделі для неперервних даних	53
Лекційне заняття № 8	53
Тема 9. Числові характеристики розподілів даних.....	57
Лекційне заняття № 9	57
ТЕМА 10. Роль граничних теорем в аналізі великих масивів даних	78
Лекційне заняття № 10	78
Лекційне заняття № 11	85

РОЗДІЛ 3. ОСНОВИ СТАТИСТИЧНОГО АНАЛІЗУ ДАНИХ	111
Тема 11. Узагальнення та представлення даних	111
Лекційне заняття № 12	111
Тема 12. Точкове оцінювання параметрів стохастичних моделей	124
Лекційне заняття № 13	124
Тема 13. Інтервальне оцінювання параметрів та довірчі інтервали.....	136
Лекційне заняття № 14	136
Тема 14. Перевірка статистичних гіпотез в аналітичних задачах	144
Лекційне заняття № 15	144
ВИКОРИСТАНІ І РЕКОМЕНДОВАНІ ЛІТЕРАТУРНІ ДЖЕРЕЛА	152
ДОДАТКИ.....	155

ВСТУП

Сучасний інформаційний світ характеризується стрімким зростанням обсягів інформації та зростаючою роллю методів інтелектуального аналізу даних у вирішенні наукових, економічних та управлінських завдань. Ефективне ухвалення рішень у реальному секторі економіки, фінансах, маркетингу та бізнесі дедалі більше спирається на дата-орієнтовані стратегії, прогнозування та статистичне оцінювання ризиків. Одним із ключових інструментів такого аналізу є стохастичні методи, які дозволяють будувати ймовірнісні моделі, оцінювати невизначеність та виявляти приховані закономірності у великих масивах даних. Застосування цих методів є фундаментом для сучасного аналітичного інструментарію та алгоритмів машинного навчання, що сприяє підвищенню точності економічних прогнозів та розробці оптимальних стратегій управління.

У багатьох сферах, зокрема в промисловому маркетингу та економіці, необхідність стохастичного аналізу зумовлена багатовимірністю та складністю реальних процесів, які генерують величезні потоки даних із численними випадковими факторами. Прогнозування коливань ринкових показників, моделювання динаміки попиту та пропозиції, глибоке дослідження поведінки споживачів та аналіз фінансових ризиків – усе це потребує глибокого розуміння ймовірнісних закономірностей та обґрунтованих методів оцінювання параметрів. Знання основних положень стохастичного моделювання дає змогу аналізувати такі процеси на основі кількісних метрик, що є необхідною складовою підготовки сучасних фахівців з маркетингу, економіки та бізнес-аналітики.

Цей навчальний посібник розроблено для надання здобувачам вищої освіти фундаментальних знань із стохастичного аналізу даних та формування практичних навичок роботи з економічними датасетами. Він охоплює основні розділи, необхідні для глибокого розуміння прикладного моделювання, а саме базові стохастичні моделі подій, застосування комбінаторики, байєсівські підходи для оновлення прогнозів, стохастичне моделювання дискретних та неперервних ознак, числові характеристики розподілів даних, а також методи узагальнення, статистичного оцінювання та перевірки гіпотез. Особливу увагу

приділено алгоритмам та статистичним методам, які використовуються в реальних економічних дослідженнях, що дозволяє студентам набувати компетентностей для застосування отриманих знань у професійній діяльності.

Метою даного посібника є формування у здобувачів фундаментальних знань та практичних навичок у сфері стохастичного аналізу для ефективного вирішення аналітичних задач, пов'язаних із дослідженням процесів, обробкою великих масивів інформації та прийняттям обґрунтованих рішень в умовах ринкової невизначеності. Своєю чергою, предметом є стохастичні процеси та закономірності, що виникають при масовій взаємодії випадкових факторів у бізнес-середовищі, а також математичні алгоритми систематизації, перевірки гіпотез та використання статистичних даних для виведення практичних висновків.

Засвоєння матеріалу посібника сприятиме формуванню важливих компетентностей, серед яких здатність до абстрактного й аналітичного мислення, вміння використовувати математичний апарат для розробки прогнозних моделей, а також розвиток навичок критичної роботи з інформацією. Розуміння принципів стохастичного моделювання дозволяє не лише оцінювати ступінь невизначеності у процесах прийняття рішень, але й обґрунтовано підходити до вибору оптимальних стратегій у висококонкурентному цифровому середовищі.

Посібник базується на програмних результатах навчання, що формуються у здобувачів під час опанування дисципліни відповідно до силабусу, і слугує важливою теоретичною базою для подальшого вивчення дисциплін, присвячених економетриці, статистиці, бізнес-аналітиці та системам штучного інтелекту. Посібник містить структуровані методологічні відомості, доповнені ілюстративними прикладами та алгоритмічними схемами, що сприяють кращому розумінню математичного інструментарію та його застосуванню в реальних умовах.

РОЗДІЛ 1. ВИПАДКОВІ ПОДІЇ ТА ЇХ МОДЕЛЮВАННЯ

ТЕМА 1. ВИПАДКОВІ ПОДІЇ ТА МІРИ НЕВИЗНАЧЕНОСТІ В БАЗОВИХ ІМОВІРНІСНИХ МОДЕЛЯХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 1

ВИПАДКОВІ І НЕВИПАДКОВІ (ДЕТЕРМІНОВАНІ) ПОДІЇ

Поняття «випадкової події» легше зрозуміти, якщо розглянути його разом з протилежним поняттям – «невипадкової події», тобто детермінованої події. А саме, детермінована подія відбувається завжди при виконанні комплексу деяких умов. Якщо посіяти зерна жита, то виросте жито, у вовків народжуються вовки, а у свиней – поросята, у людей – людські малюки. Тобто при виконанні одного того ж комплексу умов завжди відбудеться наперед відома подія.



На відміну від детермінованої події при виконанні конкретних умов випадкова подія може наступити, а може не наступити.

ПРИРОДА ВИПАДКОВОСТІ

Задаймося питанням: а в чому природа виникнення випадковості? З точки зору сучасного уявлення про оточуючий нас фізичний світ, джерело випадковості – **в квантовому характері мікросвіту**. Випадковими є перехід атомів і молекул із одних енергетичних станів в інші з випромінювання чи поглинання світла.

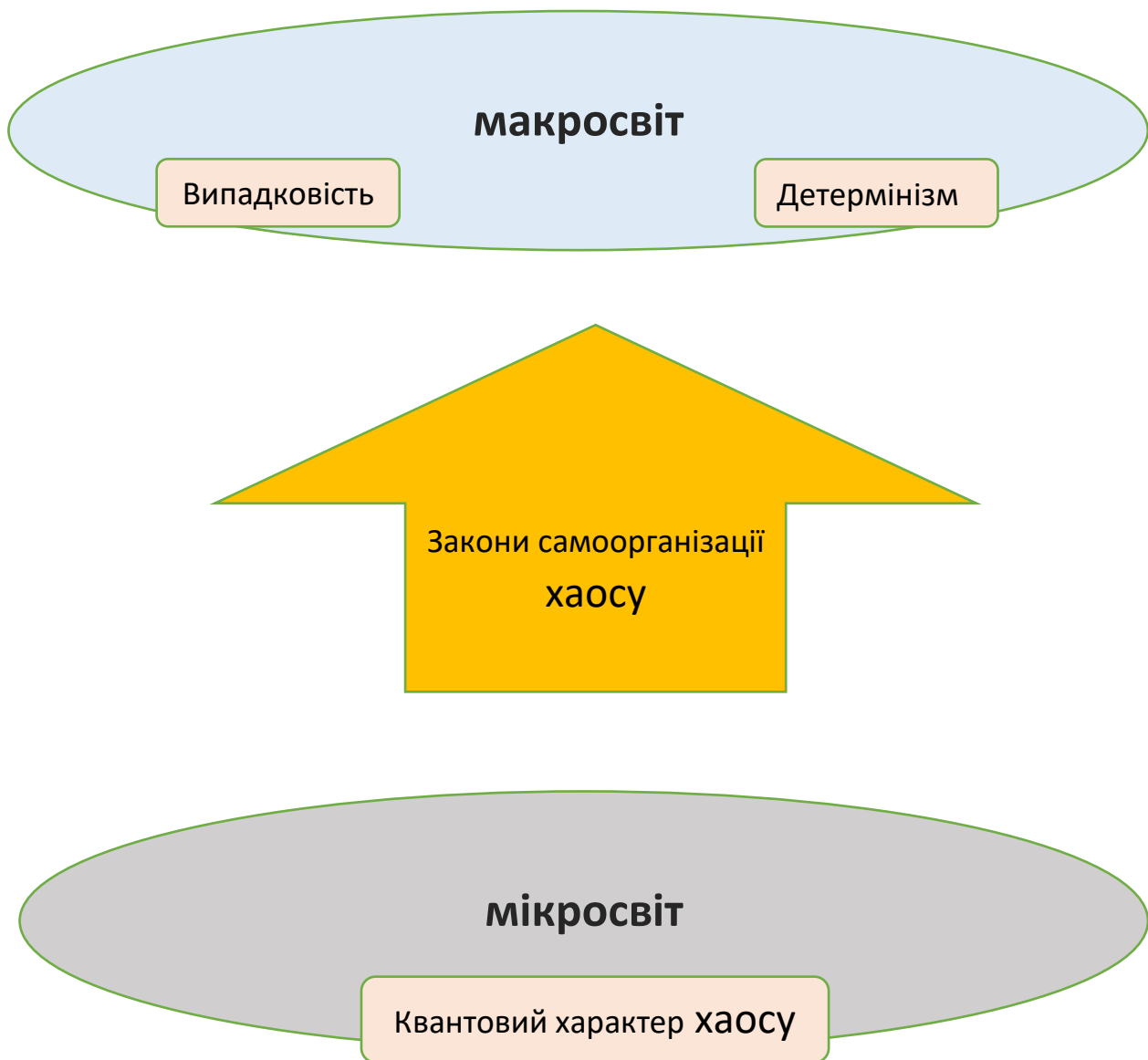
Не заглиблюючись у деталі, зазначимо, що в силу принципу невизначеності Гейзенберга неможливо з точністю меншою за певне значення визначити одночасно координати мікрочастинки та її імпульсу. Тобто у мікросвіті панує випадковість. Проявом випадковості в мікросвіті є броунівський рух.

Броунівський рух – це хаотичний тепловий рух молекул газів, рідини, плазми, що підтримується випадковими взаємодіями мікро часток. За своєю природою броунівський рух випадковий. Завдяки цьому руху відбуваються хімічні реакції (дифузія одних речовин в інші). Зокрема це забезпечує всі життєві процеси в живих організмах.

До речі, броунівський рух є гарною ілюстрацією поведінки фірми на ринку (аналог частинки – це фірма, координати частинки – є економічні параметри фірми в фазовому просторі). Велика кількість факторів впливу на фірму (процес стикання з іншими частинками) ілюструє складність прогнозування траєкторії фірми у фазовому просторі.

В оточуючому нас світі є багато явищ, які наперед ми не можемо передбачити. Так, наприклад, ми не можемо сказати чи будуть мати успіх наші дії по виведенню нового товару на ринок. А наскільки кількість товару, що знайде свого покупця, буде відрізнятися від запланованої кількості. Якщо ми виробляємо продукт, то скільки буде бракованих виробів? Ми плануємо прибуток від нашого бізнесу, але не впевненні, що його величина буде співпадати з нашим запланованим обсягом.

Джерело: Скороход А.В. Ймовірність навколо нас. 1980. Київ: «Наукова думка». 196 с.



Тому виникає питання: чи можна якимось «точними» (кількісними) методами вивчати випадковість? Здається, що кількісний підхід і випадковість виключають одне одного, несумісні з точки зору здорового глузду.

Але це не так. Виявляється, природа докільля влаштована таким чином, що нерегулярність (випадковість) окремих явищ супроводжується регулярністю поведінки їх сукупності. А наукою, що описує цей феномен, природи слугує теорія ймовірностей. Саме цією галуззю знань створена **міра випадковості – ймовірність**.

Теорія ймовірностей - математична дисципліна, об'єктом вивчення якої є випадкові події. Якщо відомі ймовірності деяких подій, то вираховуються ймовірності більш складних подій. На теорії ймовірностей ґрунтується математична статистика.

Завданням **математичної статистики** є визначення по наявному набору експериментальних даних загальних характеристик випадкових подій або явищ. В свою чергу математична статистика разом з прикладною статистикою утворюють базу аналізу даних:



За кілька десятиліть з теорії ймовірностей виділився цілий ряд самостійних напрямків: теорія випадкових процесів, теорія масового обслуговування; теорія інформації тощо.

Економіка, аналіз ринків і виробничі процеси - одна з найважливіших сфер застосування теорії ймовірності та математичної статистики. Дослідження і прогнозування економічних явищ важко собі уявити без використання методів статистичного оцінювання і перевірки гіпотез, регресійного аналізу, трендових згладжувань, економетричних моделей та інших методів, що спираються на теорію ймовірностей. З розвитком суспільства економіка все більше ускладнюється і, отже, за законами розвитку динамічних систем має посилюватися статистичний характер законів, що описують соціально-економічні явища.

Математичну основу теорії ймовірностей було реалізовано в аксіоматичній підході визначення ймовірності. Система аксіом теорії ймовірностей базується на теорії міри і була побудована відомим математиком А.М.Колмогоровим в 1933 році.

Ймовірність - це кількісна оцінка невизначеності.

Коли ми розглядаємо випадкові події, то природньо запитати, коли шанси виконатися тої чи іншої випадкової події більше, або менше.

Приклад 1. Маємо два інвестиційних проєкта. Нам треба вибрати один з них. Інтуїтивно ми розуміємо, що треба вибрати той із них, який має більші шанси на успіх. Як оцінити ці шанси? В повсякденному житті обирають той проєкт, який, як кажуть, має «більшу ймовірність» виконання.

Зробимо невеликий екскурс в історію спроб людства кількісно оцінити ймовірності випадкових подій

СТАТИСТИЧНЕ (ЧАСТОТНЕ) ВИЗНАЧЕННЯ ЙМОВІРНОСТІ

Нехай маємо стохастичний експеримент, який можливо повторити стільки раз скільки забажаємо. В ході цього експерименту розглядаємо деяку випадкову подію **A**. Повторимо цей експеримент **n** раз. Нехай

$K_n(A)$ – кількість експериментів, в яких сталася подія **A**.

Відношення $\frac{K_n(A)}{n}$ називається відносною частотою випадкової події **A** в серії стохастичних експериментів.

За допомогою цього поняття сформулюємо частотне визначення ймовірності.

$$P(A) \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{K_n(A)}{n}$$

Це означає, що для великих значень n маємо:

$$P(A) \approx \frac{K_n(A)}{n}$$

Наприклад, при великій кількості підкидань монети, відносна частота підкидань мало відрізняється від 0,5.

Експериментатор	Кількість підкидань	Кількість випадінь герба	Відносна частота
Ж. Бюффон	4040	2048	0,5080
К .Пірсон	12000	6019	0,5016
К .Пірсон	24000	12012	0,5005

Якщо при великих значеннях n відносна частота

$$\frac{K_n(A)}{n}$$

події A мало відрізняється від деякого значення P , то кажуть, що подія A стохастично стійка, а число P є ймовірністю події A .

ПОНЯТТЯ СТОХАСТИЧНОГО ЕКСПЕРИМЕНТУ. ЕЛЕМЕНТАРНА ПОДІЯ. ПРОСТІР ЕЛЕМЕНТАРНИХ ПОДІЙ. ВИПАДКОВА ПОДІЯ

Щоб почати опис випадковості розглянемо декілька понять, які лежать в основі такого опису.

Стохастичний експеримент це подія, результат якої заздалегідь не можна передбачити.

Елементарна подія – це конкретний результат здійснення стохастичного експерименту.

Простір елементарних подій – це множина усіх можливих елементарних подій.

Приклад 2. Припустимо, що підкидується монета. Елементарними подіями слугують: герб (будемо позначати Г) та решка (будемо позначати Р). Всі можливі елементарні події утворюють простір елементарних подій:

$$\Omega = \{Г, Р\}$$

Приклад 3. Підкидається гральний кубик. На верхній грані може випасти одне з шести значень. Тому простір елементарних подій має вид:

$$\Omega = \{1,2,3,4,5,6\}$$

Приклад 4. Поспіль двічі підкидається монета. Множина всіх можливих елементарних подій утворює простір елементарних подій:

$$\Omega = \{ГГ, ГР, РГ, РР\}$$

Приклад 5. Магазин працює з 8-00 до 20-00. Фіксуємо час входу до магазину конкретного покупця. Це може бути будь-який момент з часового інтервалу від 8-00 до 20-00:

$$\Omega = \{(\text{година; хвилина}): 8 \leq \text{година} \leq 20; 0 \leq \text{хвилина} < 60\}$$

Окрім елементарних подій реалії економічних відносин, менеджменту бізнесу підштовхує до розгляду більш складних подій.

Звісно випадкова подія може цікавити нас з різних причин, що пов'язані з різниці цілями. Але в жодному разі ми маємо справу з феноменом *випадкової події*. Тобто можемо казати, що (випадкова) подія відбувається, якщо реалізується одна (і тільки одна) з можливих елементарних подій. А це є підмножина простору елементарних подій. Можемо окреслити, що таке випадкова подія:

Випадкова подія – це сукупність елементарних подій, що можуть з'явитися в стохастичному експерименті.

У випадку стохастичного експерименту в прикладі 2 наведемо наступний приклад випадкової події.

Приклад 6. Підкинули гральний кубик. З якоїсь причина цікавить подія, що випаде парне число. Простір елементарних подій має вид:

$$\Omega = \{1,2,3,4,5,6\}$$

Як можна охарактеризувати ситуацію появи парного значення на гральному кубіку? Натуральне, що це можна представити як множину елементарних подій, які мають парне значення:

$$\{2,4,6\}.$$

А це є підмножиною простору елементарних подій Ω .

Формально можна визначити випадкову подію, як підмножину простору елементарних подій.

Виділяють неможливу подію

$\emptyset$ (пуста множина)

і достовірну подію

Ω (універсальна множина)

Приклад 7. Підкинули гральний кубик. Простір елементарних подій має вид:

$$\Omega = \{1,2,3,4,5,6\}$$

Розглянемо події

$$A = \{2,4\}, \quad B = \{2,4,6\}$$

Очевидно, що

$$A = \{2,4\} \subset \{2,4,6\} = B$$

Тоді A і B причинні події: якщо відбувається A , то відбувається і B , подія A є причиною події B .

КЛАСИЧНЕ ВИЗНАЧЕННЯ ЙМОВІРНОСТІ

(випадок рівноможливих наслідків)

Якщо ми проводимо стохастичний експеримент, в якому кожна елементарна подія має однакові шанси здійснитися, то маємо ситуацію класичного визначення ймовірності. Прикладами тут можуть слугувати: підкидання монети, підкидання грального кубика, тощо. Це ситуації коли кожна елементарна подія «має однакову ймовірність». Звісно, поняття «однакової ймовірності» ґрунтується тут на нашому повсякденному досвіді і фізичній інтуїції. Тим не менш, було б природньо кожній з N елементарній події в ситуації їх однакової рівно можливості надати однакову ймовірність, а саме $1/N$. Для більш складних подій ймовірність визначати з урахуванням того скільки елементах подій утворюють цю подію. Маємо класичне визначення ймовірності.

Якщо в стохастичному експерименті всі елементарні події мають однакові шанси здійснитися (однакову ймовірність), то ймовірність випадкової події **A** визначається відношенням кількості сприятливих елементарних подій **M** до кількості всіх елементарних подій **N** :

$$P(A) = \frac{M}{N}$$

Приклад 7. Підкинули гральний кубик. Простір елементарних подій має вид:

$$\Omega = \{1,2,3,4,5,6\}$$

Очевидно, що поява кожної елементарної має однакову ймовірність 1/6. Нехай нас цікавить ймовірність випадкової події, яка полягає в тому, що з'явиться парне число:

$$A = \{2,4,6\}.$$

Тобто кількість елементарних подій, що сприяють події A дорівнює 3, тоді згідно класичного визначення ймовірності

$$P(A) = \frac{3}{6} = \frac{1}{2}$$

ЙМОВІРНОСНА МОДЕЛЬ ЕКСПЕРИМЕНТА З КІНЦЕВИМ АБО ЗЛІЧЕНОЮ КІЛЬКІСТЮ ЕЛЕМЕНТАРНИХ ПОДІЙ

Розглянемо стохастичний експеримент з кінцевою або зліченою кількістю елементарних подій

$$\Omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_i, \dots\}$$

Припустимо, що кожній елементарній події ω_i надано деяку «вагу» p_i яку ми в подальшому будемо називати «ймовірністю». Припустимо також, що ці ваги задовольняють умовам:

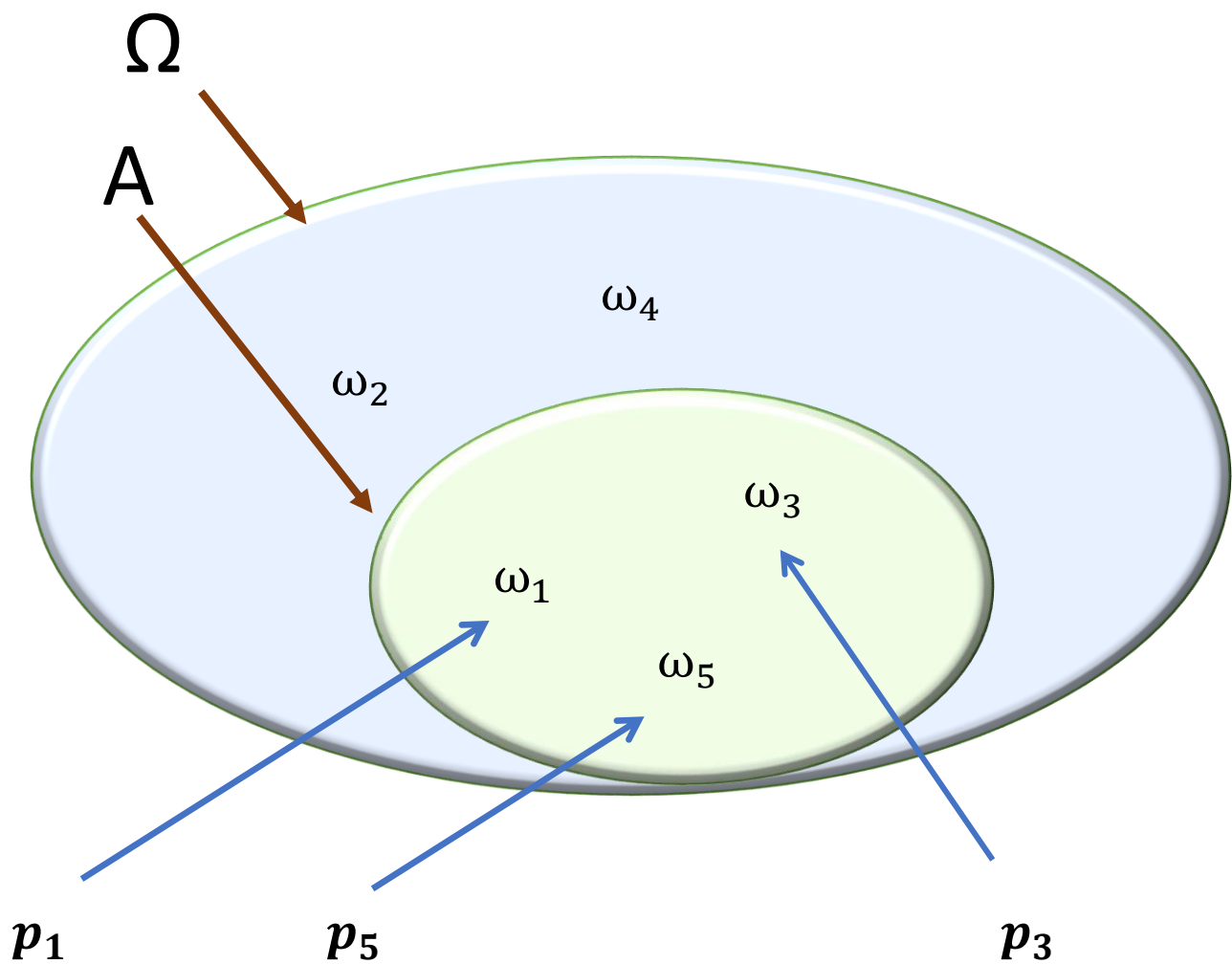
$$1) \quad p_i \geq 0. \qquad 2) \quad \sum_{i=1}^{\infty} p_i = 1$$

$$\Omega = \{ \omega_1, \omega_2, \omega_3, \dots, \omega_i, \dots \}$$

$$\begin{array}{ccccccc} \downarrow & \downarrow & \downarrow & & \downarrow \\ p_1 & p_2 & p_3 & \dots & p_i \end{array}$$

Визначення. Ймовірністю $P(A)$ випадкової події A називають суму ймовірностей елементарних подій, що утворюють подію A :

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



$$P(A) = p_1 + p_3 + p_5$$

ЙМОВІРНОСНА МОДЕЛЬ ЕКСПЕРИМЕНТА З КІНЦЕВОЮ АБО ЗЛІЧЕНОЮ КІЛЬКІСТЮ ЕЛЕМЕНТАРНИХ ПОДІЙ

1. $\Omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_i, \dots\}$ - простір елементарних подій
2. $\Omega \supset A$ – випадкові події
3. $P(\cdot)$ – ймовірність (ймовірнісна міра):

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

а) $p_i \geq 0$.

б) $\sum_{i=1}^{\infty} p_i = 1$

ЙМОВІРНІСТНА МОДЕЛЬ СТОХАСТИЧНОГО ЕКСПЕРИМЕНТА З НЕЗЛІЧЕНИМ ЧИСЛОМ ЕЛЕМЕНТАРНИХ ПОДІЙ. ГЕОМЕТРИЧНА ЙМОВІРНІСТЬ

Застосувати підхід побудови ймовірнісної моделі стохастичного експерименту зі скінченною або зліченою множиною елементарних подій не виходить. У випадку, коли простір елементарних подій має геометричну природу визначення ймовірності може бути задано наступним чином. Тут випадкова подія **A** являє собою геометричну фігуру:

- 1) підмножину на числовій прямій, яка має довжину;
- 2) підмножину на площині, яка має площу;
- 3) підмножину у просторі, яка має об'єм.

Для уніфікації запису будемо застосовувати позначення

***mes* (A)**

mes від **measure** (англ. Міра)

ВИЗНАЧЕННЯ ГЕОМЕТРИЧНОЇ ЙМОВІРНОСТІ

Якщо Ω — деяка множина геометричної природи, що є простором елементарних подій, **A** — випадкова подія (тобто її підмножина). Тоді ймовірність події **A** визначається наступним чином:

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

Задача «про зустріч». Два студента вирішили зустрітися. Кожен з них не зміг гарантувати точний час зустрічі. Домовилися таким чином: протягом часового інтервалу $[0, T]$ кожен з'являється в обумовленому місці. Оскільки у кожного напружений графік діяльності, то домовилися, що час очікування один одного не буде перевищувати τ ($\tau < T$). Питання: з якою ймовірністю вони зустрінуться?

Рішення. Всі можливі ситуації – це набір пар:

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

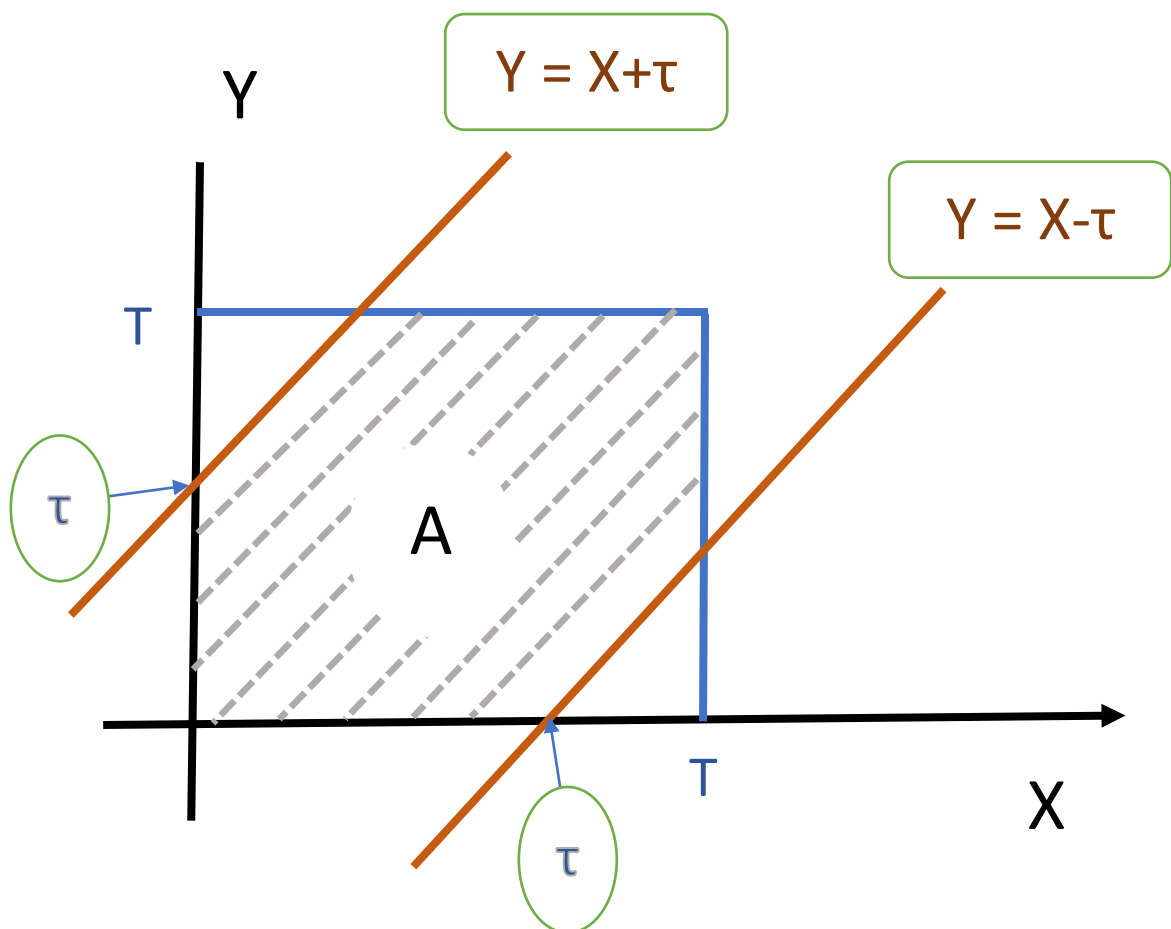
де x – час приходу на зустріч першого студента, а y – час приходу на зустріч другого студента. Нехай A – випадкова подія, яка полягає в тому, що зустріч в відбудеться:

$$A = \{(x; y): |x - y| \leq \tau\}$$

Зауважено, що

$$|x - y| \leq \tau \Leftrightarrow x - \tau \leq y \leq x + \tau$$

Геометрично це можна візуалізувати наступним чином:



Проведемо обчислення

$$mes(\Omega) = T \cdot T = T^2$$

$$mes(A) = T^2 - (T - \tau)^2$$

Згідно визначення геометричної ймовірності:

$$P(A) = \frac{mes(A)}{mes(\Omega)} = \frac{T^2 - (T - \tau)^2}{T^2}$$

У окремому випадку, якщо

$$T = 1 \text{ год}, \quad \tau = 15 \text{ хв.} = \frac{1}{4} \text{ год.}$$

Маємо

$$P(A) = \frac{mes(A)}{mes(\Omega)} = \frac{1^2 - (1 - 1/4)^2}{1^2} = 1 - \left(\frac{3}{4}\right)^2 = \frac{7}{16}$$

АКСІОМАТИЧНЕ ВИЗНАЧЕННЯ ЙМОВІРНОСТІ (КОНЕПТУАЛЬНИЙ ПОГЛЯД)

Будуємо ймовірнісний простір $(\Omega, \mathcal{F}, P)$, що складається з трьох елементів.

I. Ω – простір елементарних подій;

II. $\mathcal{F}$ – множина випадкових подій. Це деяка множина підмножин простору елементарних подій:

$$A \subset \Omega$$

(але не всі підмножини – це випадкові події !!!)

III. P – функція, яка кожній випадковій події A ставить у відповідність число від 0 до 1 і задовольняє таким умовам:

1) $0 \leq P(A)$

2) $P(\Omega) = 1$

3) A_i – послідовність несумісних подій:

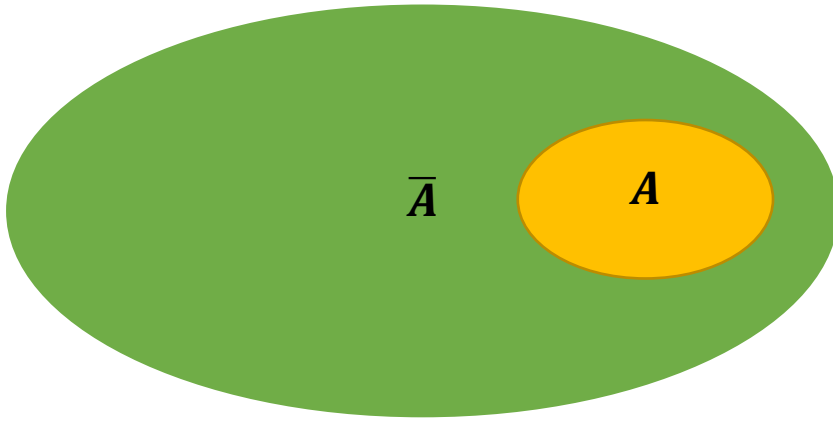
$$A_i \cap A_j = \emptyset, \quad i \neq j$$

Тоді

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

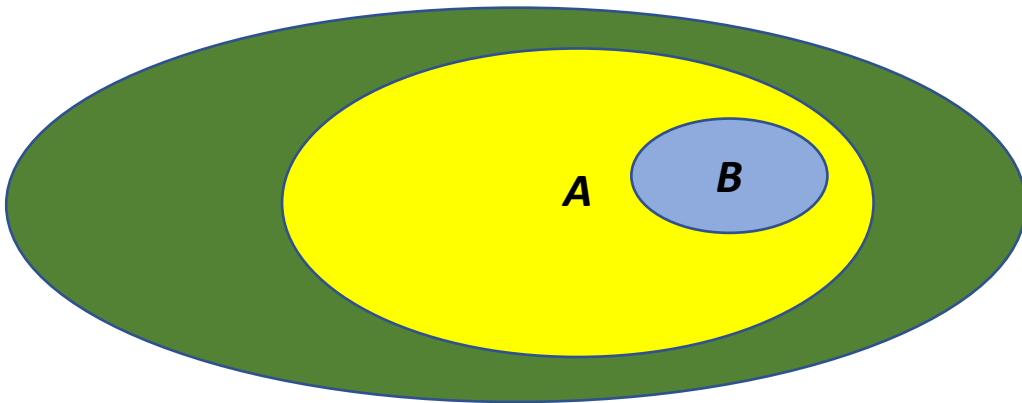
ВЛАСТИВОСТІ ЙМОВІРНОСТІ

1) $P(\bar{A}) = 1 - P(A)$



2) Якщо

$$A \subset B, \text{ то } P(A) \leq P(B)$$



3) $P(\emptyset) = 0$

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

ТЕМА 2. ЗАСТОСУВАННЯ КОМБІНАТОРИКИ ДЛЯ ОБЧИСЛЕННЯ ЙМОВІРНОСТЕЙ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 2

Для розрахунку числа елементарних подій в класичній схемі можна використовувати формули комбінаторики.

Перестановки:

$$P_n = n!$$

Кількість перестановок – кількість способів, якими n предметів можна розмістити на n місць.

Розміщення:

$$A_n^k = \frac{n!}{(n-k)!}$$

Кількість розміщень – кількість способів, якими n предметів можна розмістити на k місць.

Комбінації:

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

Кількість комбінацій – кількість способів, якими можна вибрати k предметів з n без урахування порядку вибору.

Розміщення з повтореннями:

$$\bar{A}_n^k = n^k$$

Комбінації з повтореннями:

$$\bar{C}_n^k = C_{n+k-1}^k$$

Для використання формул комбінаторики у розрахунках кількості сприятливих деякій події наслідків, необхідно відповісти на два питання:

1. Чи має значення порядок вибору елементів?

2. Чи можливі повторення елементів (схема з поверненням або без повернення)?

Особливості експерименту		Повторення	
		Немає	Є
Порядок вибору	Неважливий	$C_n^k = \frac{n!}{k!(n-k)!}$	$\bar{C}_n^k = C_{n+k-1}^k$
	Важливий	$A_n^k = \frac{n!}{(n-k)!}$	$\bar{A}_n^k = n^k$

ТЕМА 3. УМОВНА ЙМОВІРНІСТЬ ТА НЕЗАЛЕЖНІСТЬ ПОДІЙ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 3

УМОВНА ЙМОВІРНІСТЬ

Приклад 1. Нехай маємо стохастичний експеримент, який полягає в підкиданні два рази грального кубика. Простір елементарних подій:

$$\Omega = \{(i; j): i = 1, 2, 3, 4, 5, 6; j = 1, 2, 3, 4, 5, 6\}$$

Розглянемо дві події:

$$A = \{\text{перший раз з'явиться одиниця}\} = \\ = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\}$$

$$B = \{\text{сума очок, що випала не перебільшує 3}\} = \\ = \{(1, 1), (1, 2), (2, 1)\}.$$

Відомо, що B відбулося. Нехай $P(A/B)$ – ймовірність події A при умові, що подія B відбулася. Обчислимо ймовірності:

$$P(A) = \frac{6}{36} = \frac{1}{6}; \quad P(B) = \frac{3}{36} = \frac{1}{12}; \quad P(A \cap B) = \frac{2}{36} = \frac{1}{18}$$

Визначення. A і B - випадкові події ($P(B)$). Умовною ймовірністю події A за умови, що подія B відбулася називають

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

ВЛАСТИВОСТІ УМОВНОЇ ЙМОВІРНОСТІ

$$1) \quad 0 \leq P(A/B)$$

$$2) \quad P(\Omega / B) = 1$$

$$3) \quad P(B / B) = 1$$

4) A_i – послідовність несумісних подій:

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

тоді

$$P\left(\bigcup_i A_i / B\right) = \sum_{i=1}^{\infty} P(A_i / B)$$

Доведення. За визначенням умовної ймовірності

$$\begin{aligned} P\left(\bigcup_i A_i / B\right) &= \frac{P((\bigcup_i A_i) \cap B)}{P(B)} = \frac{\sum_i P(A_i \cap B)}{P(B)} = \\ &= \sum_i \frac{P(A_i \cap B)}{P(B)} = \sum_i P(A_i / B) \end{aligned}$$

ФОРМУЛА МНОЖЕННЯ.

Нехай A і B - випадкові події ($P(A) \neq 0, P(B) \neq 0$). Тоді

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

НЕЗАЛЕЖНІ ПОДІЇ

ВИЗНАЧЕННЯ. Нехай A і B – випадкові події. Події A і B називаються незалежними, якщо

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

ТЕОРЕМА. Нехай A і B – випадкові події,

$$P(A) > 0, \quad P(B) > 0$$

Тоді

$$A \text{ і } B \text{ - незалежні} \iff \begin{cases} P(A) = P(A/B) \\ P(B) = P(B/A) \end{cases} \quad (*)$$

Доведення. Необхідність. Нехай A і B - незалежні події. Згідно означення незалежності випадкових подій

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

Звідси за означенням умовної ймовірності

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

та

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

Достатність. Нехай виконано (*). Тоді, згідно з означенням умовної ймовірності та (*) маємо

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

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

Тобто

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

За означенням A і B - незалежні події.

ТЕМА 4. ФОРМУЛА ПОВНОЇ ЙМОВІРНОСТІ У МОДЕЛЮВАННІ СКЛАДНИХ ПОДІЙ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 4

Щоб навести формулу повної ймовірності знадобиться поняття повної групи подій

Визначення. Кажуть, що набір випадкових подій

$$H_1, H_2, \dots, H_n$$

утворюють *повну групу* подій, якщо:

$$1) H_1 \cup H_2 \cup \dots \cup H_n = \Omega$$

$$2) A_i \cap A_j = \emptyset, \quad (i \neq j) \quad (\text{тобто попарно несумісні})$$

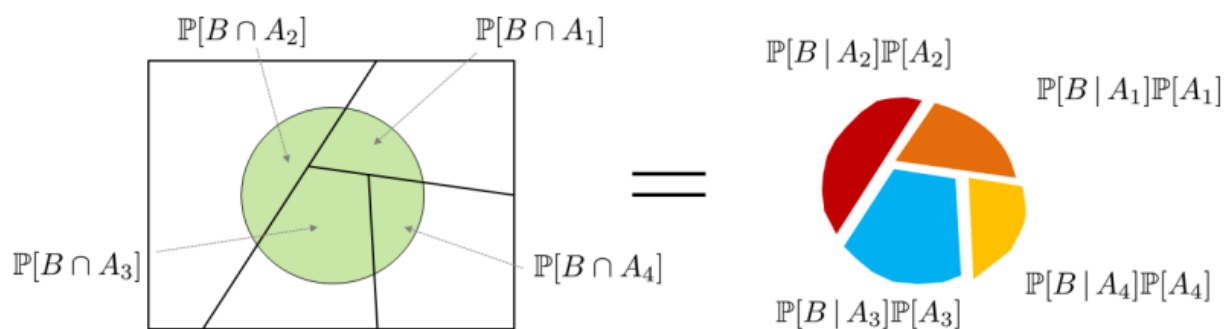
ТЕОРЕМА (ФОРМУЛА ПОВНОЇ ЙМОВІРНОСТІ). Нехай A - випадкова подія,

$H_1, H_2, \dots, H_n$ - повна група подій ($P(H_i) \neq 0$). Тоді має місце формула повної ймовірності:

$$P(A) = P(H_1) \cdot P(A/H_1) + \dots + P(H_n) \cdot P(A/H_n)$$

Доведення.

$$P(A) = \sum_{i=1}^{\infty} P(H_i \cap A) = \sum_{i=1}^{\infty} P(H_i) \cdot P(A/H_i)$$



Задача про «щасливий білет». Всього маємо N білетів. Серед них n «щасливих». Питання: у кого більше шансів вибрати «щасливий» білет: хто відійде першим, або хто відійде другим?

Розглянемо дві випадкові події:

$A = \{\text{перший взяв "щасливий" білет}\}$

$B = \{\text{другий взяв "щасливий" білет}\}$

$$P(A) \gtrless P(B) ?$$

Розглянемо гіпотези:

$$H_1 = \{\text{перший взяв "щасливий" білет}\}$$

$$H_2 = \{\text{перший не взяв "щасливий" білет}\}$$

H_1 і H_2 – повна група подій,

$$P(H_1) = \frac{n}{N} = P(A), \quad P(H_2) = \frac{N - n}{N}$$

Умовні ймовірності:

$$P(B/H_1) = \frac{n - 1}{N - 1}, \quad P(B/H_2) = \frac{n}{N - 1}$$

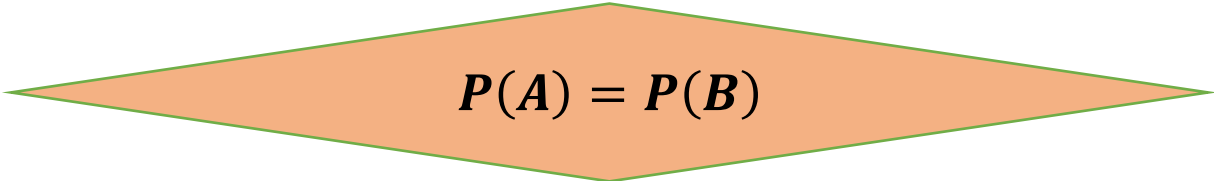
За формулою повної ймовірності

$$\begin{aligned} P(B) &= P(H_1) \cdot P(B/H_1) + P(H_2) \cdot P(B/H_2) = \\ &= \frac{n}{N} \cdot \frac{n - 1}{N - 1} + \frac{N - n}{N} \cdot \frac{n}{N - 1} = \frac{n}{N} \end{aligned}$$

Оскільки

$$P(A) = \frac{n}{N},$$

то


$$P(A) = P(B)$$

ТЕМА 5. ФОРМУЛА БАЙЄСА ТА ПЕРЕОЦІНКА ЙМОВІРНОСТЕЙ НА ОСНОВІ НОВИХ ДАНИХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 5

ТЕОРЕМА. Нехай $H_1, H_2, \dots, H_n$ - повна група подій, ($P(H_i) \neq 0$),

A - випадкова подія ($P(A) \neq 0$). Тоді мають місце формули Байєса

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(H_1) \cdot P(A/H_1) + \dots + P(H_n) \cdot P(A/H_n)}$$

$$i = 1, 2, \dots, n$$

Доведення.

З формул множення

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

Враховуючи формулу повної ймовірності в знаменнику правої частини, отримаємо, що треба довести.

ЗАСТОСУВАННЯ ФОРМУЛ БАЙЄСА

Загальний алгоритм застосування формул Байєса при вирішенні практичних проблем полягає в наступному. Вважаємо, що подія A здійснюється в умовах, щодо яких можна зробити припущення у виконанні гіпотетичних умов (гіпотез) $H_1, H_2, \dots, H_n$. З якихось міркувань відомі

ймовірності цих гіпотез $P(H_1), P(H_2), \dots, P(H_n)$. Це ймовірності, які відомі до спостереження (*апріорні ймовірності = до експериментальні*). Також є можливість дізнатися про умовні ймовірності

$$P(A/H_1), P(A/H_2), \dots, P(A/H_n)$$

Припустимо, що подія **A** здійснилася. Це потребує переоцінку гіпотез, тобто виявлення ймовірності гіпотез при умові додаткової інформації (здійснення події **A**):

$$P(H_1/A), P(H_2/A), \dots, P(H_n/A)$$

Це можливо за допомогою формул Байеса. Оскільки відповідні ймовірності розглядаються після того як відомо що подія **A** здійснилася, то такі ймовірності називаються *апостеріорними ймовірностями (=після експериментальні)*.

Діагностування хвороби. Хвороба зустрічається один випадок хвороби на тисячу людей (1:1000). Тест показав, що людина хвора. Надійність тесту 99%, тобто у 99% випадків виявляє хворобу.

З якою ймовірністю людина, що має позитивний результат тесту дійсно хвора?

Розв'язок. Нехай

$$H_1 = \{\text{людина хвора}\}$$

$$H_2 = \{\text{людина здорова}\}$$

Очевидно H_1, H_2 - повна група подій і

$$P(H_1) = \frac{1}{1000} = 0.001$$

$$P(H_2) = \frac{999}{1000} = 0.999$$

Позначимо випадкову подію

$$A = \{\text{тест вказує на хворобу}\}$$

Оскільки тест виявляє хворобу у 99%, то

$$P(A/H_1) = 0.99, \quad P(A/H_2) = 0.01$$

За формулами Байєса

$$\begin{aligned} P(H_1/A) &= \frac{P(H_1) \cdot P(A/H_1)}{P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2)} = \\ &= \frac{0,001 \cdot 0,99}{0,001 \cdot 0,99 + 0,999 \cdot 0.01} = 0.0902 \approx 0.09 \end{aligned}$$

Таким чином ймовірність того, що отримавши позитивний тест, людина хвора дорівнює 9%.

Якщо ця ж людина зробить повторно тест і він знову буде позитивним. З якою ймовірністю людина, що має два незалежних позитивних результатів тестів, хвора?

Нехай

$H_3 = \{\text{людина хвора і має позитивний перший тест}\}$

$H_4 = \{\text{людина здорова і має позитивний перший тест}\}$

Можемо вважати, що

$$P(H_3) = P(H_1/A) = 0.0902$$

Позначимо випадкову подію

$B = \{2 \text{ тести позитивних тести}\}$

За формулами Байєса

$$\begin{aligned} P(H_3/B) &= \frac{P(H_3) \cdot P(B/H_3)}{P(H_3) \cdot P(B/H_3) + P(H_4) \cdot P(B/H_4)} = \\ &= \frac{0,0902 \cdot 0,99}{0,0902 \cdot 0,99 + 0,9098 \cdot 0,01} = 0.9075 \approx 0.91 \end{aligned}$$

Таким чином при двох позитивних тестах ймовірність, що людина хвора дорівнює 91%.

ТЕМА 6. МОДЕЛЮВАННЯ ПОСЛІДОВНИХ ВИПРОБУВАНЬ ЗА СХЕМОЮ БЕРНУЛЛІ ЛЕКЦІЙНЕ ЗАНЯТТЯ № 6

ВИЗНАЧЕННЯ. Проводиться скінчена кількість n незалежних випробування, в кожному з яких може відбуватися деяка подія (будемо називати її «успіхом» і позначати «У»), або протилежна їй подія (будемо називати її «неуспіхом» і позначати «Н»). Така послідовність випробувань називається *схемою Бернуллі*.

Приклад 1. Нехай маємо урну з білими та чорними шарами. Навмання вибираємо з урни один шар, після чого повертаємо його до урни. Далі обираємо другий шар і також повертаємо до урни. Цю процедуру повторюємо третій раз. Таким чином маємо $n=3$ незалежних випробувань. В якості «успіху» розглядаємо появу білого шару. Тоді «неуспіхом» буде поява чорного шару. Вважаємо, що частка білих шарів у урні p а частка чорних, відповідно, $q=1-p$. Процедури виймання *незалежні*, то маємо схему Бернуллі. Множина всіх можливих серій випробувань, тобто простором елементарних подій буде множина:

$$\Omega = \{УУУ, УУН, УНУ, УНН, НУУ, НУН, ННУ, ННН\}$$

Якщо

$$P(Y) = p ,$$

то

$$P(H) = 1 - p \equiv q$$

Оскільки підкидання незалежні, то ймовірність елементарних подій:

$$P(УУУ) = p^3$$

$$P(УУН) = p^2 \cdot q$$

$$P(УНУ) = p^2 \cdot q$$

$$P(УНН) = p \cdot q^2$$

$$P(НУУ) = p^2 \cdot q$$

$$P(НУН) = p \cdot q^2$$

$$P(ННУ) = p \cdot q^2$$

$$P(ННН) = q^3$$

Розглянемо випадкові події

$$A_0 = \{\text{не відбудеться жодного } Y\}$$

$$A_1 = \{\text{відбудеться один "У" }\}$$

$$A_2 = \{\text{відбудеться два "У" }\}$$

$$A_3 = \{\text{відбудеться три "У" }\}$$

Очевидно

$$P(A_0) = P(\text{HHH}) = q^3$$

$$P(A_1) = P(\text{УНН, НУН, ННУ}) = 3 \cdot p \cdot q^2$$

$$P(A_2) = P(\text{УУН, УНУ, НУУ}) = 3 \cdot p^2 \cdot q$$

$$P(A_3) = P(\text{УУУ}) = p^3$$

Зауважимо, що кількість елементів множин A_0, A_1, A_2, A_3 є відповідно сполучення $C_3^0 = 1, C_3^1 = 3, C_3^2 = 3, C_3^3 = 1$, оскільки означають скільки підмножин відповідно розміром 0,1,2,3 можна відібрати з трьохелементної множини.

Розглянемо загальний випадок. В схемі Бернуллі з n незалежними випробуваннями позначимо

$$A_k = \{\text{в серії випробувань відбудеться } k \text{ "У" } \}$$

$$(i = 0, 1, 2, \dots, n)$$

Оскільки випробування незалежні, то ймовірність елементарних подій виду

$$\omega = \underbrace{\{Y, Y, H, \dots \dots \dots, H, Y, H\}}_{k \text{ разів } Y \text{ та } n-k \text{ разів } H}$$

дорівнює

$$p^k \cdot q^{n-k}$$

Кількість таких елементарних подій – це кількість k -елементних підмножин n -елементарної множини, тобто сполучення з n по k :

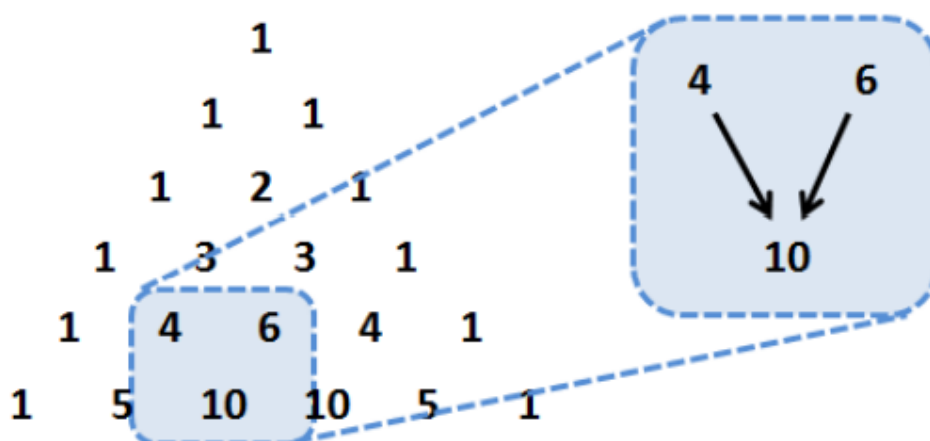
$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{1 \cdot 2 \cdot 3 \cdot \dots \cdot n}{(1 \cdot 2 \cdot \dots \cdot k) \cdot (1 \cdot 2 \cdot \dots \cdot (n-k))}$$

Отже доведено твердження.

ТЕОРЕМА. Якщо проводиться n незалежних випробувань в кожному з яких може відбутися подія («успіх») з ймовірністю p або не відбутися («неуспіх») з ймовірністю $1 - p \equiv q$. Тоді ймовірність того, що в цій серії випробувань відбудеться k «успіхів»:

$$P(\text{відбудеться } k \text{ "успіхів"}) = C_n^k \cdot p^k \cdot q^{n-k}$$

Зауважимо, що значення сполучень з n по k можна знайти з трикутника Паскаля (n – номер рядка, k – номер числа в рядку).



Трикутник Паскаля, де кожне число отримується як сума двох чисел над ним.

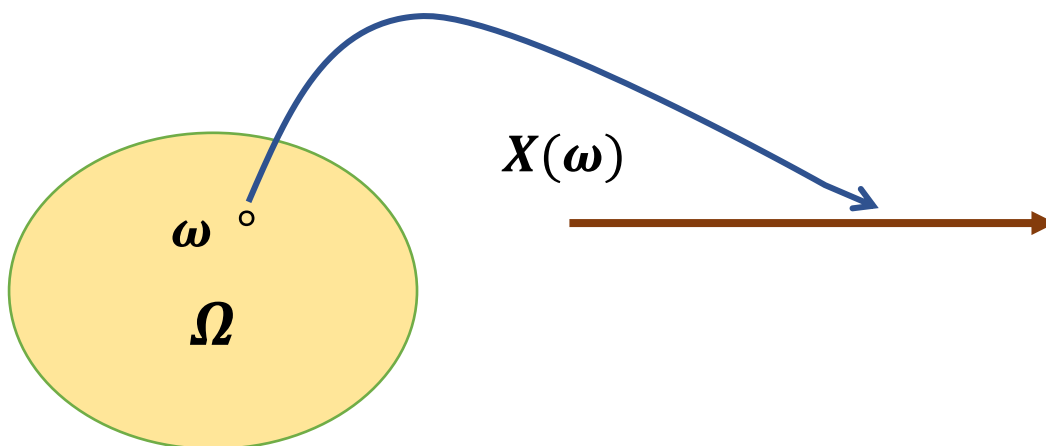
РОЗДІЛ 2. ЙМОВІРНІСНІ МОДЕЛІ ДАНИХ (ВИПАДКОВІ ВЕЛИЧИНИ)

ТЕМА 7. СТОХАСТИЧНІ МОДЕЛІ ДЛЯ ДИСКРЕТНИХ ДАНИХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 7

ВИПАДКОВІ ВЕЛИЧИНИ

Випадкова величина – це функція, яка визначена на просторі елементарних подій.



Приклади, щодо економічних систем: ринковий попит на товар, кількість рентабельних підприємств галузі або території, дохід/прибуток підприємства в майбутньому тощо.

Наприклад, розглядаємо стохастичний експеримент, що полягає в підкиданні два рази монети.

Приклад 1. Поспіть двічі підкидається монета. Множина всіх можливих елементарних подій утворює простір елементарних подій:

$$\Omega = \{\Gamma\Gamma, \Gamma P, P\Gamma, PP\}$$

Нехай випадкова величина $X(\omega)$ – це кількість гербів, що випала, тобто

$$X(\Gamma\Gamma) = 2, \quad X(\Gamma P) = 1$$

$$X(P\Gamma) = 1 \quad X(PP) = 0$$

Випадкова величина приймає три значення?: 0,1,2. Можемо підрахувати ймовірності, з якими випадкова величина приймає ці значення.

$$P(X = 0) = P(PP) = 1/4$$

$$P(X = 1) = P(\Gamma P, P\Gamma) = 1/4 + 1/4 = 1/2$$

$$P(X = 2) = P(\Gamma\Gamma) = 1/4$$

Визначення. Якщо $(\Omega, \mathcal{F}, P)$ - ймовірнісний простір.

(Ω -простір елементарних подій, $\mathcal{F}$ -множина випадкових подій,

P -ймовірність). Тоді функція

$$X(\omega): \Omega \longrightarrow \mathbb{R}$$

Називається **випадковою величиною (random variable)**, якщо для
любого $t \in \mathbb{R}$

$$\{\omega \mid X(\omega) < t\} \in \mathcal{F}$$

тобто є випадковою подією.

Можна довести, що множини виду

$$\bigcup_i \{a_i \leq X(\omega) < b_i\}$$

та

$$\bigcap_i \{a_i \leq X(\omega) < b_i\}$$

є випадковими подіями, тобто для них існує ймовірність.

ДИСКРЕТНІ ВИПАДКОВІ ВЕЛИЧИНИ

Визначення. Кажуть, що випадкова величина має дискретний розподіл, якщо множина її значень скінчена або злічена.

В прикладі 1 наведена випадкова величина, яка дорівнює числу гербів, що з'явилися в стохастичному експерименті з двократним підкиданням монети. Вона приймає три значення: 0, 1, 2. Згідно означення вона має дискретний розподіл.

Приклад 2. Підкинули гральний кубик. Простір елементарних подій має вид:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Очевидно, що поява кожної елементарної події має однакову ймовірність $1/6$.

Розглянемо випадкову величину $X(\omega)$, що дорівнює числу крапок на верхній грані. Тоді вона приймає скінчену множину

$$\begin{aligned} P(X = 1) &= P(X = 2) = P(X = 3) = \\ &= P(X = 4) = P(X = 5) = P(X = 6) = 1/6 \end{aligned}$$

Заради компактності запису ймовірності, з якими випадкова величина приймає свої значення, можна записати в наступному виді:

<i>X</i>	1	2	3	4	5	6
<i>P</i>	1/6	1/6	1/6	1/6	1/6	1/6

Таку таблицю називають **рядом розподілу**.

Для прикладу 1 ряд розподілу має вид:

<i>X</i>	0	1	2
<i>P</i>	1/4	1/2	1/4

Дихотомічний розподіл (розподіл Бернуллі)

Випадкова величина, що приймає два значення (наприклад 0 і 1) має дихотомічний розподіл. Її ряд розподілу:

<i>X</i>	0	1
<i>P</i>	<i>q</i>	<i>p</i>

Тут **$q = 1 - p$**

Біноміальний розподіл

Розглянемо схему Бернуллі, тобто серію з n незалежних випробувань, в кожному з яких може відбуватися «Успіх» («У») з ймовірності p , або протилежна подія «Неуспіхом» («Н») з ймовірністю $q=1-p$. Говорять, що випадкова величина X має біноміальний розподіл, якщо вона дорівнює кількості «У» в цій серії випробувань:

$$P(X = k) = C_n^k \cdot p^k \cdot q^{n-k}$$

k – кількість «У»,

$$C_n^k = \frac{n!}{k!(n-k)!}$$

Позначення. Якщо біноміально розподілена випадкова величина X розглядається в n незалежних випробуваннях з ймовірністю успіху в кожному випробуванні p , то це позначається

$$X \sim B(n, p).$$

Скільки логічних аналізаторів запланувати для виробництва. (Siegel, 2016). Складність виробництва логічних аналізаторів не позбавлена недоліків. Попередня статистика виробництва показує, що 97% вироблених аналізаторів задовольняють встановленим стандартам. Треба відвантажити споживачу 17 аналізаторів, що відповідають стандартам. Питання: скільки треба виробити аналізаторів, щоб мати розумні гарантії відвантажити 17 добре працюючих логічних аналізаторів?

Якщо буде вироблено 17 аналізаторів яка ймовірність належним чином виконати замовлення?

Зауважимо, що ймовірність «успіху» (тобто вироблений аналізатор задовольняє стандартам) дорівнює

$$P = 0.97$$

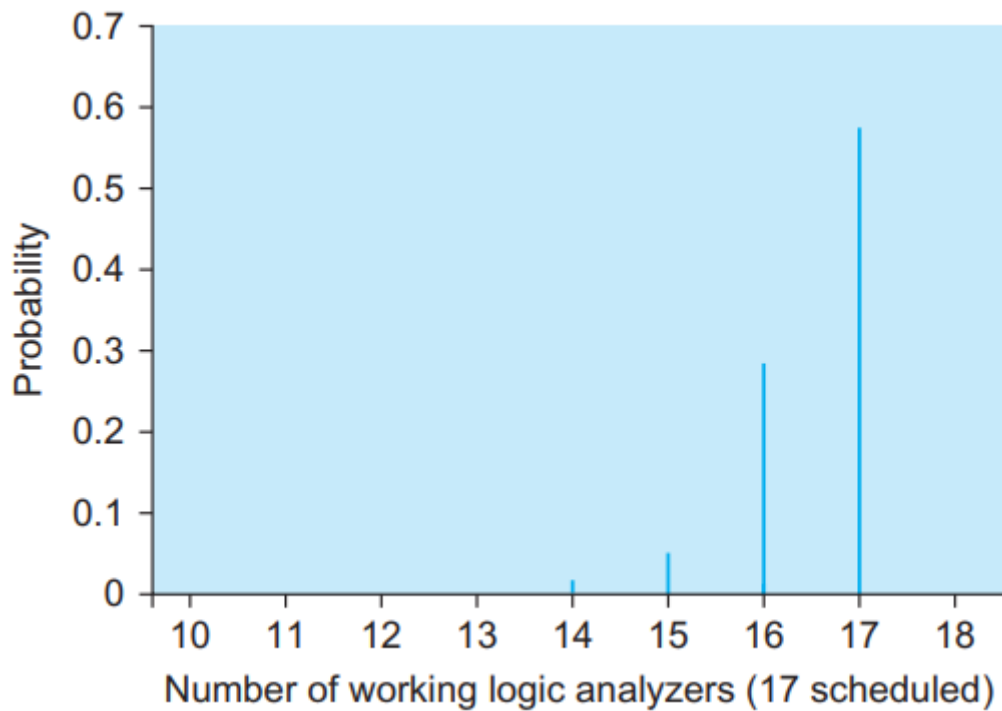
Розглянемо ситуацію, коли вироблено 17 аналізаторів. Якщо X позначає кількість аналізаторів, які можна відвантажити споживачу, то

$$X \sim B(17, 0.97).$$

Тоді

$$\begin{aligned} P(X = 17) &= C_{17}^{17} \cdot (0.97)^{17} \cdot (0.03)^0 = \\ &= 1 \cdot 0.595826 \cdot 1 = 0.595826 \approx 0.596 \end{aligned}$$

Таким чином при виробництві 17 аналізаторів ймовірність логічного аналізатора задовольняти встановленим стандартам дорівнює 0,596 (59,6%).
--



Розглянемо ситуацію, коли вироблено 18 аналізаторів.

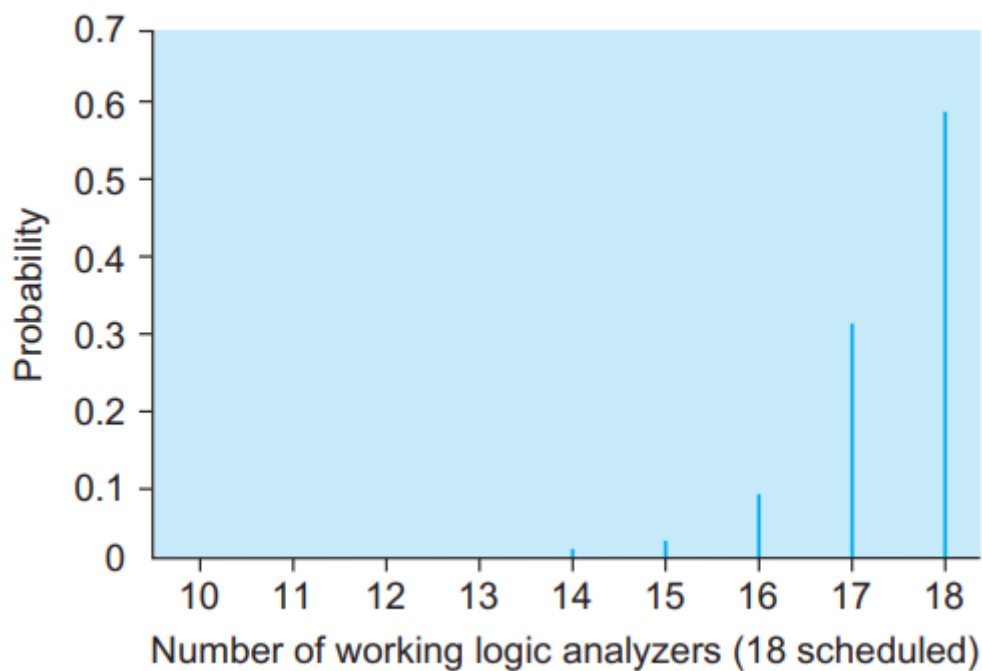
Якщо Y позначає кількість аналізаторів, які можна відвантажити споживачу, то

$$Y \sim B(18, 0.97).$$

Тоді

$$\begin{aligned}
 P(\text{можна відгрузити } 17 \text{ аналізаторів}) &= \\
 &= P(Y = 17) + P(Y = 18) = \\
 &= C_{18}^{17} \cdot (0.97)^{17} \cdot (0.03)^1 + C_{18}^{18} \cdot (0.97)^{18} \cdot (0.03)^0 = \\
 &= 18 \cdot 0.595826 \cdot 0.03 + 1 \cdot 0.577951 \cdot 1 \\
 &= 0.321746 + 0.577951 = 0.899697 \approx 0.900
 \end{aligned}$$

Таким чином при виробництві 18 аналізаторів ймовірність логічного аналізатора задовольняти встановленим стандартам дорівнює 0,900 (90%).



Розглянемо ситуацію, коли вироблено 19 аналізаторів.

Якщо **Z** позначає кількість аналізаторів, які можна відвантажити споживачу, то

$$Z \sim B(19, 0.97).$$

Тоді

$$\begin{aligned}
 P(\text{можна відгрузити } 17 \text{ аналізаторів}) &= \\
 &= P(Z = 17) + P(Z = 18) + P(Z = 19) = \\
 &= C_{19}^{17} \cdot (0.97)^{17} \cdot (0.03)^2 + C_{19}^{18} \cdot (0.97)^{18} \cdot (0.03)^1 + \\
 &+ C_{19}^{19} \cdot (0.97)^{19} \cdot (0.03)^0 = \frac{18 \cdot 19}{1 \cdot 2} \cdot (0.97)^{17} \cdot (0.03)^2 + \\
 &+ 19 \cdot 0.577951 \cdot 0.03 + 1 \cdot 0.560613 \cdot 1 = 0.981742 \approx 0.982
 \end{aligned}$$

Таким чином при виробництві 19 аналізаторів ймовірність логічного аналізатора задовольняти встановленим стандартам дорівнює 0,982 (98,2%).

Геометричний розподіл

Розглянемо серію незалежних випробувань, в кожному з яких може відбуватися «Успіх» («У») з ймовірності p , або протилежна подія «Неуспіх» («Н») з ймовірністю $q=1-p$. Проводимо випробування до першої появи «У». Простір елементарних подій:

$$\Omega = \{У, НУ, ННУ, НННУ, \dots\}.$$

З умов незалежності випробувань

$$P(У) = p, \quad P(НУ) = q \cdot p,$$

$$P(ННУ) = q^2 \cdot p, \dots, P(\underbrace{НН \dots Н}_{k-1} У) = q^{k-1} \cdot p$$

Говорять, що випадкова величина X має «У»: **геометричний розподіл**, якщо вона дорівнює кількості випробувань до першої появи «У»

$$P(X = k) = q^{k-1} \cdot p$$

Мартингальна стратегія (стратегія подвоєння). Граємо в «орлянку». Підкидуємо монету. Якщо випав «герб», то отримаємо нагороду в **1 (грошову одиницю=ГО)**, а, якщо «решітка», то програємо. Стратегія «мартингал» полягає у тому, що, подвоюючи ставку до того моменту, коли отримаємо виграш. **Питання:** якщо ми маємо 100 го (фінансові, адміністративні обмеження), то дотримуючись стратегії «мартингала», з якою ймовірністю отримаємо результат?

Якщо на першому етапі маємо «герба», то зупиняємося з отриманням 1 (1 го). Далі діємо до моменту випадіння «герба». Якщо програли, то подвоюємо ставку тощо.

Маємо обмеження

$$1, 2, 4, 8, 16, 32, 64 < 100$$

$$1, 2, 4, 8, 16, 32, 64, 128 > 100$$

Якщо X - кількості випробувань до першої появи «У» («Герба» з ймовірністю $p = \frac{1}{2}$) урахуванням цих обмежень

$$X \leq 7.$$

Тоді ймовірність виграти дорівнює

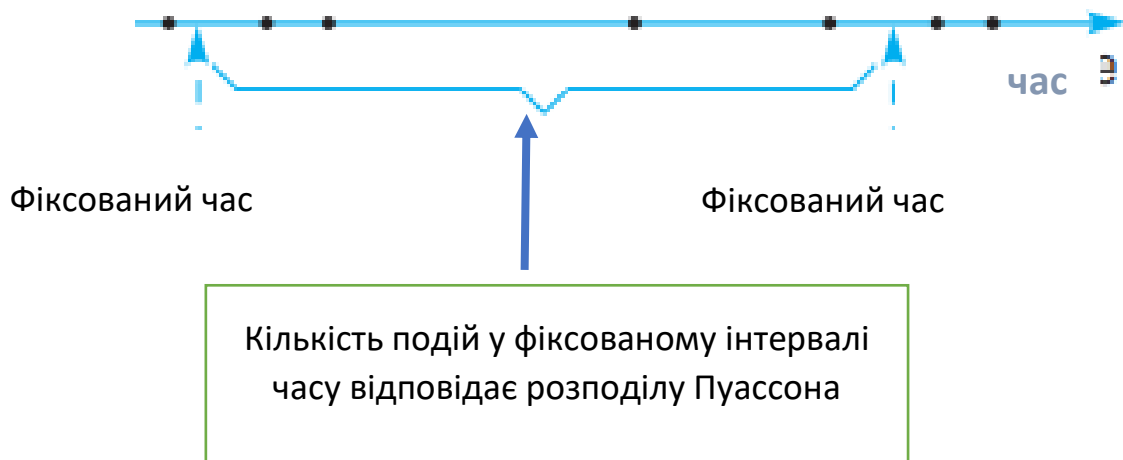
$$\begin{aligned} P(\text{виграємо маючи } 100 \text{ го.}) &= \\ &= P(X = 1) + P(X = 2) + \dots + P(X = 7) = \\ &= q^0 \cdot p + q^1 \cdot p + \dots + q^6 \cdot p = \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{64} = \frac{63}{64} \approx 0.984 \end{aligned}$$

Розподіл Пуассона

Випадкова величина X має розподіл Пуассона з параметром $\lambda > 0$ якщо

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Розподіл Пуассона грає значну роль в теорії масового обслуговування, теорії надійності, тощо.



Закон рідкісних подій

Вважаємо, що:

- 1) Події здійснюються в різні моменти (не можуть здійснитися одночасно);
- 2) Послідовність подій – послідовність незалежний подій;
- 3) В середньому кількість подій постійна.



Джерело: <https://cutt.ly/VrevpbgL>

«1» - подія відбувається з ймовірністю P

«0» - подія не відбувається з ймовірністю q ($q=1-p$);

Теорема Пуассона. Нехай $X(n) \sim B(n, p(n))$, тобто має біноміальний розподіл в n незалежних випробувань з ймовірністю «успіху» $p(n)$. До того ж $n \cdot p(n) \rightarrow \lambda > 0$. Тоді для любого $k = 0, 1, 2, \dots$ має місце

$$P(X(n) = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Визначення. Кажуть, що випадкова величина X має неперервний розподіл, якщо

$$F(t) = \int_{-\infty}^t p(u) du$$

функція $p(u)$ називається **щільністю розподілу** випадкової величини.



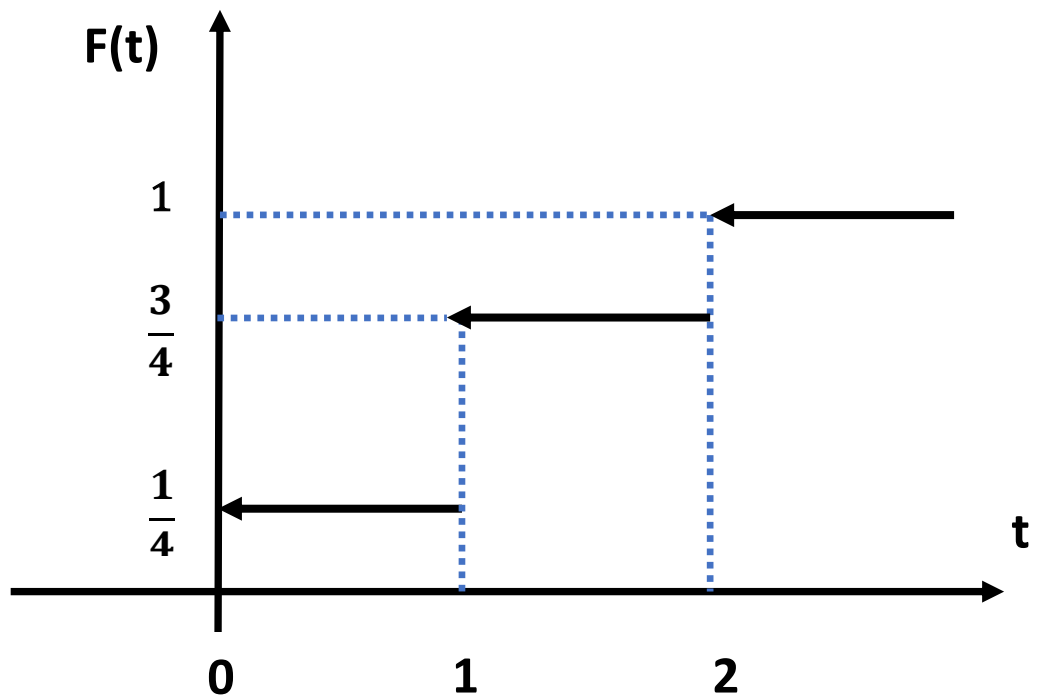
Приклад 1. Поспіль двічі підкидається монета. Простір елементарних подій:

$$\Omega = \{ГГ, ГР, РГ, РР\}$$

Нехай випадкова величина $X(\omega)$ – це кількість гербів, що випала.

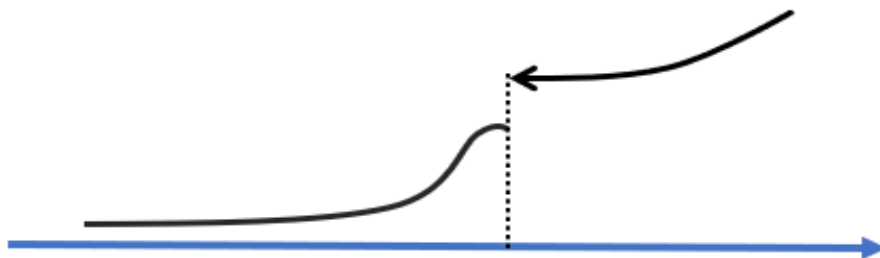
X	0	1	2
P	1/4	1/2	1/4

Побудуємо функцію розподілу:



Властивості функції розподілу

- 1) $0 \leq F(t) \leq 1$
- 2) $F(t)$ - неспадна функція (зростає або стала)
- 3) $F(t)$ – неперервна зліва



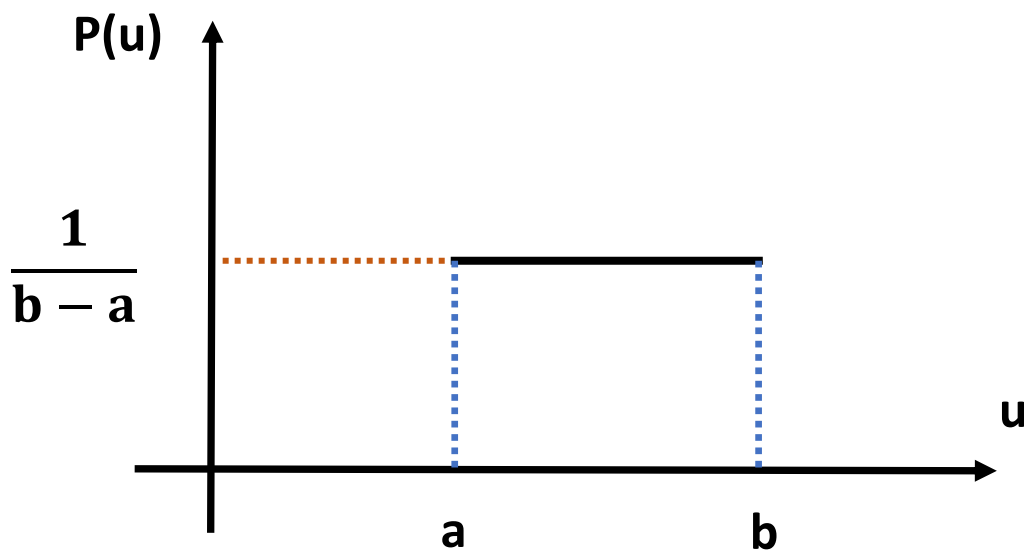
ТЕМА 8. СТОХАСТИЧНІ МОДЕЛІ ДЛЯ НЕПЕРЕРВНИХ ДАНИХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 8

Рівномірний розподіл

Випадкова величина X має рівномірний на відрізку $[a, b]$ розподіл, якщо

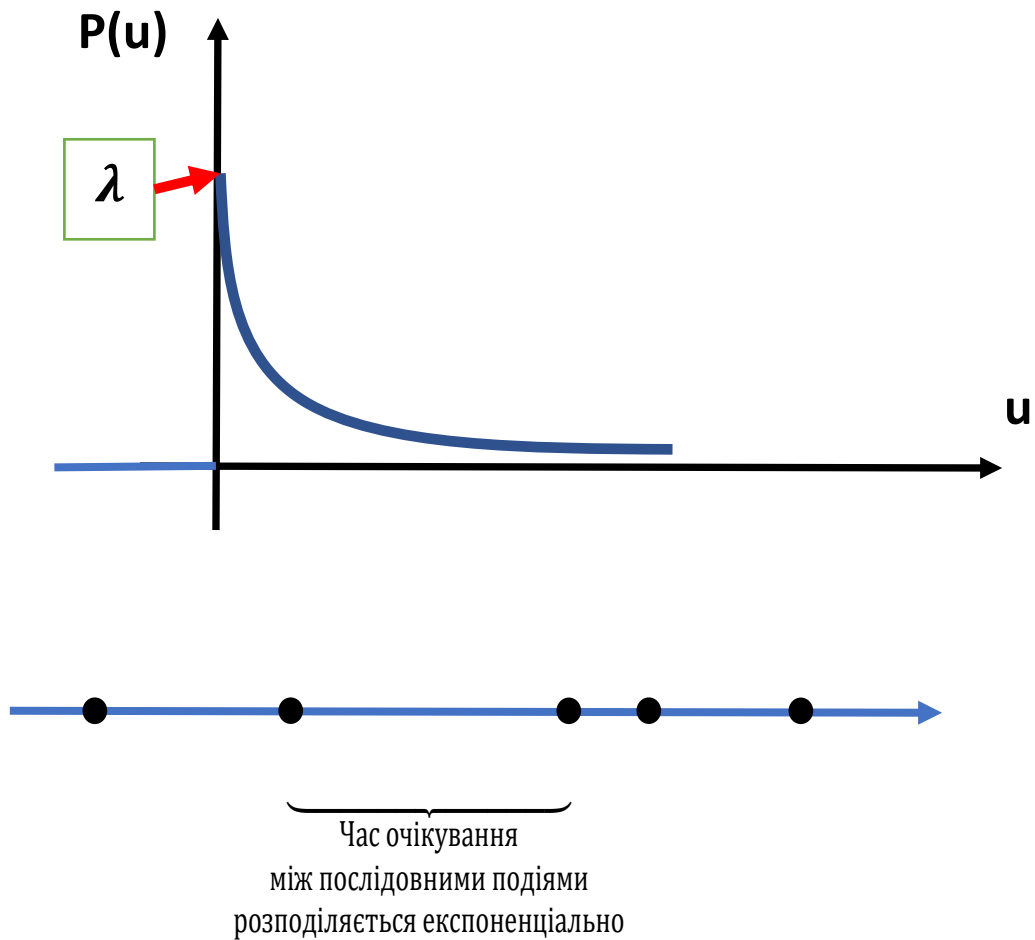
$$p(u) = \begin{cases} \frac{1}{b-a}, & u \in [a, b] \\ 0, & u \notin [a, b] \end{cases}$$



Експоненціальний розподіл

Кажуть, що випадкова величина X має експоненціальний розподіл з параметром $\lambda > 0$, якщо її щільність розподілу має вид

$$p(u) = \begin{cases} \lambda \cdot e^{-\lambda u}, & u \geq 0 \\ 0, & u < 0 \end{cases}$$

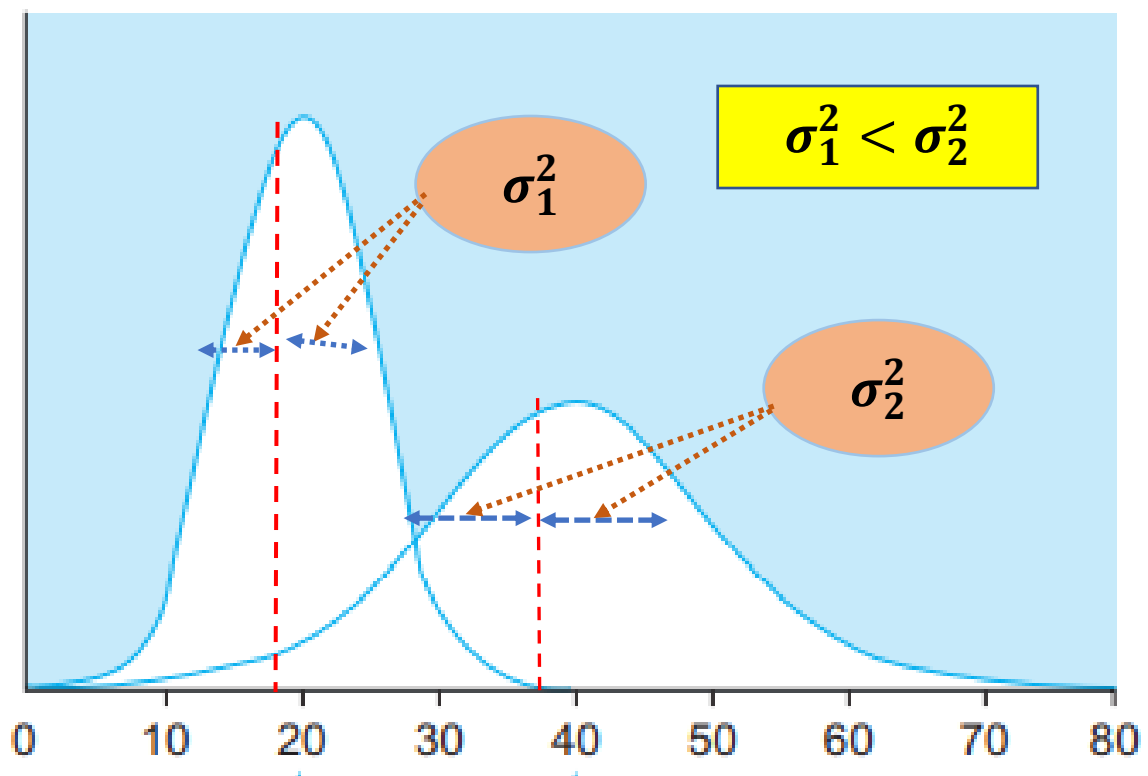
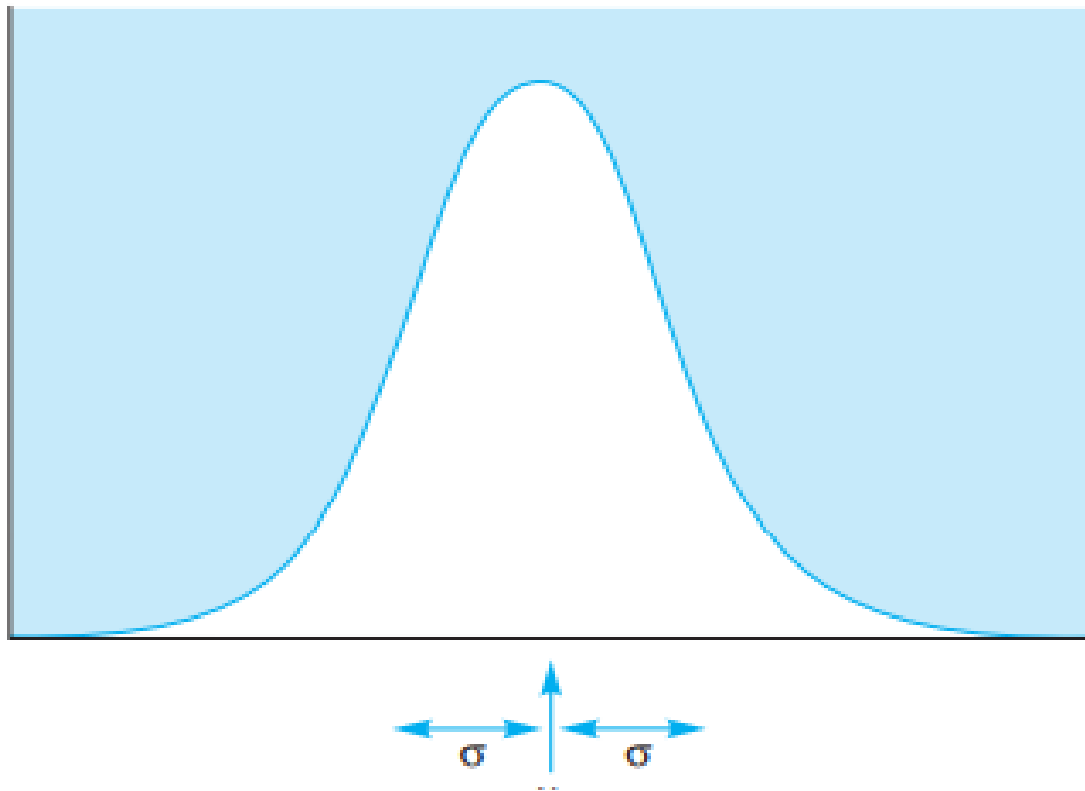


Нормальний розподіл (розподіл Гаусса)

Нормальним розподілом $N(a, \sigma^2)$ з параметрами a і σ^2 називається розподіл з щільністю розподілу

$$p(u) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp \left\{ -\frac{1}{2} \cdot \frac{(u - a)^2}{\sigma^2} \right\}$$

Щільність нормального розподілу з $a=0$ та $\sigma^2 = 1$.



Властивості щільності розподілу

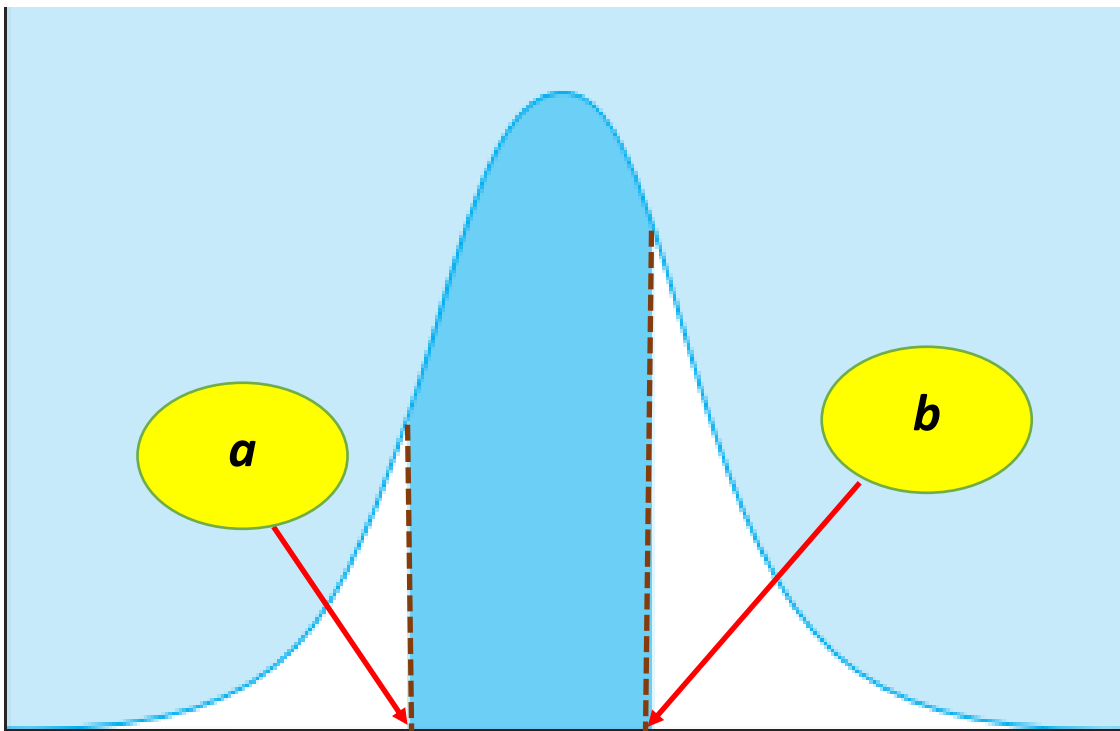
1)

$$\int_{-\infty}^{+\infty} p(u) du = 1$$

2) Якщо в точці t функція $F(t)$ - неперервна, то

$$\frac{dF(t)}{dt} = p(t)$$

3) $P(a \leq X < b) = \int_a^b p(u) du$



ТЕМА 9. ЧИСЛОВІ ХАРАКТЕРИСТИКИ РОЗПОДІЛІВ ДАНИХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 9

Математичне сподівання

Визначення. Математичним сподіванням випадкової величини X називається:

а) Якщо X дискретна величина з рядом розподілу

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

то

$$EX = x_1 p_1 + x_2 p_2 + \dots + x_n p_n + \dots$$

б) якщо X неперервна величина з щільністю $p(u)$, то

$$EX = \int_{-\infty}^{+\infty} u p(u) du$$

Властивості математичного сподівання

1. Математичне сподівання постійної дорівнює самій постійної:

$$E(C) = C.$$

Доведення. Випадкова величина X приймає тільки одне значення C з імовірністю $p = 1$, то

$$E(C) = C \cdot 1 = C.$$

2. Постійний множник можна виносить за знак математичного очікування:

$$E(CX) = C EX.$$

Доведення. (у випадку дискретної випадкової величини) Якщо випадкова величина X задана рядом розподілу

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

Тоді ряд розподілу CX має вигляд

CX	Cx_1	Cx_2	$\dots$	Cx_n	$\dots$
$\mathbb{P}$	p_1	p_2	$\dots$	p_n	$\dots$

звідси

$$\begin{aligned} ECX &= Cx_1p_1 + Cx_2p_2 + \dots + Cx_np_n + \dots = \\ &= C(x_1p_1 + x_2p_2 + \dots + x_np_n + \dots) = CEX \end{aligned}$$

3. Математичне сподівання суми двох випадкових величин дорівнює сумі математичних сподівань доданків:

$$E(X + Y) = EX + EY$$

Доведення (у випадку дискретної випадкової величини). Нехай маємо ряд розподілу випадкової величини **X** :

X	x_1	x_2	$\dots$	x_n	$\dots$
$\mathbb{P}$	p_1^x	p_2^x	$\dots$	p_n^x	$\dots$

А ряд розподілу випадкової величини **Y** має вигляд:

Y	y_1	y_2	$\dots$	y_n	$\dots$
$\mathbb{P}$	p_1^y	p_2^y	$\dots$	p_n^y	$\dots$

Обчислимо **$E(X + Y)$** :

$$\begin{aligned}
 E(X + Y) &= \sum_n \sum_k (x_n + y_k) P(X = x_n, Y = y_k) = \\
 &= \sum_n x_n \sum_k P(X = x_n, Y = y_k) + \\
 &+ \sum_k y_k \sum_n P(X = x_n, Y = y_k) = \\
 &= \sum_n x_n p_n^x + \sum_k y_k p_k^y = EX + EY
 \end{aligned}$$

Приклад 1. Нехай X – дихотомічно розподілена випадкова величина:

X	0	1
$\mathbb{P}$	$q = 1 - p$	p

Тоді

$$EX = 0 \cdot q + 1 \cdot p = p$$

Приклад 2. Нехай X – біноміальна випадкова величина:

$$P(X = k) = C_n^k p^k q^{n-k} = \frac{n!}{k! (n-k)!} p^k q^{n-k}$$

$$k=0,1,2,\dots,n$$

Нагадаємо, що розглядається схема Бернуллі, тобто n незалежних експериментів, у кожного з яких може статися «успіх» («У») з ймовірністю p і «неуспіх» («Н») з ймовірністю $q = 1-p$. Випадкова величина X дорівнює кількості «успіхів» («У») в послідовності n незалежних випробувань.

Математичне сподівання дорівнює:

$$EX = \sum_{k=0}^n k \cdot \frac{n!}{k! (n-k)!} p^k q^{n-k}$$

Технічно обчислити нелегко. Зробимо інакше. А саме представимо випадкову величину X як суму індикаторів (i – номер випробувань).

$$Y_i = \begin{cases} 1, & \text{якщо здійснився "У"}; \\ 0, & \text{якщо здійснився "Н"} \end{cases}$$

Зауважено, що відповідно до результату прикладу 1, маємо $EY_i = p$.

Тоді випадкову величину X можна представити у вигляді:

$$X = Y_1 + Y_2 + \dots + Y_n$$

Відповідно до властивості 3) математичних сподівань, маємо

$$EX = E(Y_1 + Y_2 + \dots + Y_n) =$$

$$EY_1 + EY_2 + \dots + EY_n = p + p + \dots + p = pn$$

Приклад 3. Нехай X – геометрично розподілена випадкова величина:

$$P(X = k) = p \cdot q^{k-1}, \quad k = 1, 2, 3, \dots$$

Нагадаємо, що X – кількість експериментів до першого успіху («У»), ймовірність «У» - p , а неуспіху («Н») дорівнює: $q = 1 - p$.

Тоді математичне сподівання дорівнює

$$\begin{aligned} EX &= \sum_{k=1}^{+\infty} k \cdot p \cdot q^{k-1} = p \cdot \sum_{k=1}^{+\infty} (q^k)' = \\ &= p \left(\sum_{k=1}^{+\infty} q^k \right)' = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p} \end{aligned}$$

Приклад 4. Нехай X – має розподіл Пуассону: з параметром $\lambda > 0$:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Тоді, згідно з визначенням математичного сподівання маємо

$$\begin{aligned} EX &= \sum_{k=0}^{+\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=1}^{+\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \\ &= \lambda \cdot \sum_{k=1}^{+\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot \underbrace{\sum_{m=0}^{+\infty} \frac{\lambda^m}{m!}}_{=1} = \lambda \end{aligned}$$

Приклад 5. Нехай X – рівномірно розподілена на відрітку $[a, b]$ випадкова величина, тобто щільність має вигляд:

$$p(u) = \begin{cases} \frac{1}{b-a}, & \text{якщо } u \in [a, b] \\ 0, & \text{якщо } u \notin [a, b] \end{cases}$$

Тоді, згідно з визначенням математичного сподівання маємо

$$\begin{aligned} EX &= \int_{-\infty}^{+\infty} u p(u) du = \\ &= \int_a^b u p(u) du = \frac{1}{b-a} \int_a^b u du = \frac{1}{2(b-a)} u^2 = \\ &= \frac{1}{2(b-a)} u^2 \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \end{aligned}$$

Приклад 6. Нехай X – експоненціальна випадкова величина с параметром λ ,

Тобто її щільність має вигляд:

$$p(u) = \begin{cases} \lambda e^{-\lambda u}, & \text{якщо } u \geq 0 \\ 0, & \text{якщо } u < 0 \end{cases}$$

Тоді, згідно з визначенням математичного сподівання маємо

$$\begin{aligned} EX &= \int_{-\infty}^{+\infty} u p(u) du = \int_0^{+\infty} u \cdot \lambda e^{-\lambda u} du = \\ &= - \int_0^{+\infty} u \cdot e^{-\lambda u} d(-\lambda u) = - \int_0^{+\infty} u d(e^{-\lambda u}) = \\ &= - u e^{-\lambda u} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda u} du = \\ &= - \frac{1}{\lambda} \int_0^{+\infty} e^{-\lambda u} d(-\lambda u) = - \frac{1}{\lambda} \cdot e^{-\lambda u} \Big|_0^{+\infty} = \frac{1}{\lambda} \end{aligned}$$

Приклад 7. Нехай X – нормально розподілена випадкова величина с параметрами a та σ^2 . Тобто щільність розподілу має вигляд:

$$p(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-a)^2}{2\sigma^2}}$$

Тоді

$$EX = \int_{-\infty}^{+\infty} u p(u) du = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} u e^{-\frac{(u-a)^2}{2\sigma^2}} du. \quad (*)$$

Зробимо заміну

$$t = \frac{u-a}{\sigma}.$$

Звідси

$$u = \sigma \cdot t + a; du = \sigma \cdot dt.$$

Продовжимо співвідношення (*):

$$\begin{aligned} (*) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (t\sigma + a) e^{-\frac{t^2}{2}} dt = \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t \cdot e^{-\frac{t^2}{2}} dt + a \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt \end{aligned}$$

Перший інтеграл взятий по симетричному інтервалу від непарної функції дорівнює нулю. А інтеграл дорівнює одиниці оскільки – це інтеграл від щільності розподілу стандартної нормальної випадкової величини $N(0,1)$:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$$

Отже виявилось, що параметр a в нормальному розподілі $N(a, \sigma^2)$ – це математичне сподівання випадкової величини.

$$EX = a$$

ТЕОРЕМА (про математичне сподівання функції від випадкової величини).

Нехай X – випадкова величина, а $\varphi(\cdot)$ – деяка функція. Тоді

a) Якщо X дискретна величина з рядом розподілу

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

то

$$E\varphi(X) = \varphi(x_1)p_1 + \varphi(x_2)p_2 + \dots + \varphi(x_n)p_n + \dots$$

b) якщо X неперервна величина з щільністю $p(u)$, то

$$E\varphi(X) = \int_{-\infty}^{+\infty} \varphi(u) p(u) du$$

Доведення. (у випадку дискретної випадкової величини). Якщо випадкова величина X задана рядом розподілу

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

Тоді ряд розподілу $\varphi(X)$ має вигляд

$\varphi(X)$	$\varphi(x_1)$	$\varphi(x_2)$	$\dots$	$\varphi(x_n)$	$\dots$
$\mathbb{P}$	p_1	p_2	$\dots$	p_n	$\dots$

За означенням математичного сподівання

$$E\varphi(X) = \varphi(x_1)p_1 + \varphi(x_2)p_2 + \dots + \varphi(x_n)p_n + \dots$$

ДИСПЕРСІЯ ВИПАДКОВОЇ ВЕЛИЧИНИ

Визначення. Нехай X – випадкова величина, $a = EX$. Тоді

$$DX \stackrel{\text{def}}{=} E(X - a)^2$$

називається **дисперсією** випадкової величини X .

Визначення. Нехай X – випадкова величина. **Середнім квадратичним відхиленням** називається величина

$$\sigma \stackrel{\text{def}}{=} \sqrt{DX}$$

Властивості дисперсії

1. Якщо $X = C \equiv \text{Const}$, то

$$D(X) = 0$$

2. Якщо $C \equiv \text{Const}$, то

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

3. Якщо $C \equiv \text{Const}$, то

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

4. Має місце наступна формула

$$D(X) = EX^2 - (EX)^2$$

Доведення

$$\begin{aligned} DX &\stackrel{\text{def}}{=} E(X - a)^2 = E(X^2 - 2Xa + a^2) = \\ &= EX^2 - 2aEX + a^2 = \\ &= EX^2 - 2a^2 + a^2 = EX^2 - a^2 = EX^2 - (EX)^2 \end{aligned}$$

Приклад 1. Нехай X – дихотомічно розподілена випадкова величина:

X	0	1
$\mathbb{P}$	$q = 1 - p$	p

Тоді згідно властивості 4)

$$\begin{aligned} DX &\stackrel{\text{def}}{=} E(X - p)^2 = EX^2 - p^2 = \\ &= 0 \cdot q + 1 \cdot p - p^2 = \\ &= p - p^2 = p(1 - p) = pq \end{aligned}$$

Приклад 2. Нехай X – геометрично розподілена випадкова величина:

$$P(X = k) = p \cdot q^{k-1}, \quad k = 1, 2, 3, \dots$$

Нагадаємо, що X – кількість експериментів до першого успіху («У»), ймовірність «У» - p , а неуспіху («Н») дорівнює: $q = 1$

Скористуємося властивістю 4). Для цього обчислимо

$$\begin{aligned} EX^2 &= \sum_{k=1}^{+\infty} k^2 \cdot p \cdot q^{k-1} = \\ &= \sum_{k=1}^{+\infty} k \cdot (k - 1 + 1) \cdot p \cdot q^{k-1} = \\ &= \sum_{k=1}^{+\infty} k \cdot (k - 1) \cdot p \cdot q^{k-1} + \underbrace{\sum_{k=1}^{+\infty} k \cdot p \cdot q^{k-1}}_{EX} \\ &= p \cdot q \cdot \sum_{m=0}^{+\infty} \underbrace{m \cdot (m - 1) \cdot q^{m-2}}_{(q^m)''} + \frac{1}{p} = \\ &= p \cdot q \cdot \left(\sum_{m=0}^{+\infty} q^m \right)'' + \frac{1}{p} = p \cdot q \cdot \left(\frac{1}{1 - q} \right)'' + \frac{1}{p} \\ &= \frac{2 \cdot p \cdot q}{(1 - q)^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p} \end{aligned}$$

Приклад 3. Нехай X – має розподіл Пуассону: з параметром $\lambda > 0$:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Тоді, згідно з теоремою про математичне сподівання функції від випадкової величини маємо

$$\begin{aligned} EX^2 &= \sum_{k=0}^{+\infty} k^2 \cdot e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{+\infty} k \cdot (k-1+1) \cdot \frac{\lambda^k}{k!} = \\ &= e^{-\lambda} \sum_{k=2}^{+\infty} k \cdot (k-1) \frac{\lambda^k}{k!} + \underbrace{\sum_{k=0}^{+\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!}}_{EX} = \end{aligned}$$

$$\begin{aligned} &= e^{-\lambda} \cdot \lambda^2 \cdot \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda = \\ &= \lambda^2 \cdot e^{-\lambda} \cdot \underbrace{\sum_{m=0}^{+\infty} \frac{\lambda^m}{m!}}_{=1} + \lambda = \lambda^2 + \lambda \end{aligned}$$

Враховуючи властивість 4), маємо

$$DX = EX^2 - (EX)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Приклад 4. Нехай X – рівномірно розподілена на відрізку $[a, b]$ випадкова величина, тобто щільність має вигляд:

$$p(u) = \begin{cases} \frac{1}{b-a}, & u \in [a, b] \\ 0, & u \notin [a, b] \end{cases}$$

Тоді, згідно з визначенням дисперсії та теореми про математичне сподівання функції випадкової величини маємо (позначимо $m = EX$)

$$\begin{aligned} DX &\stackrel{\text{def}}{=} E(X - m)^2 = \int_{-\infty}^{+\infty} (u - m)^2 p(u) du = \\ &= \int_a^b (u - m)^2 p(u) du = \\ &= \frac{1}{(b-a)} \int_a^b (u - m)^2 du = \\ &= \frac{1}{(b-a)} \frac{(u - m)^3}{3} \Big|_a^b = \frac{(b-a)^2}{12} \end{aligned}$$

оскільки $m = \frac{a+b}{2}$.

Приклад 5. Нехай X – експоненціальна випадкова величина с параметром

$\lambda > 0$, тобто її щільність має вигляд:

$$p(u) = \begin{cases} \lambda e^{-\lambda u}, & \text{якщо } u \geq 0 \\ 0, & \text{якщо } u < 0 \end{cases}$$

Згідно теореми про математичне сподівання функції випадкової величини маємо

$$EX^2 = \int_{-\infty}^{+\infty} u^2 p(u) du = \int_0^{+\infty} u^2 \cdot \lambda e^{-\lambda u} du =$$

$$= - \int_0^{+\infty} u^2 \cdot e^{-\lambda u} d(-\lambda u) = - \int_0^{+\infty} u^2 d(e^{-\lambda u}) =$$

$$= - u^2 e^{-\lambda u} \Big|_0^{+\infty} + 2 \cdot \int_0^{+\infty} u \cdot e^{-\lambda u} du =$$

$$= \frac{2}{\lambda} \cdot \underbrace{\int_0^{+\infty} u \cdot \lambda e^{-\lambda u} du}_{EX = \frac{1}{\lambda}} = \frac{2}{\lambda^2}$$

Враховуючи властивість 4), маємо

$$DX = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Приклад 6. Нехай X – нормально розподілена випадкова величина с параметрами a та σ^2 . Тобто щільність розподілу має вигляд:

$$p(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-a)^2}{2\sigma^2}}$$

Було доведено, що

$$EX = a$$

Обчислимо дисперсію

$$\begin{aligned} DX &\stackrel{\text{def}}{=} E(X - a)^2 = \int_{-\infty}^{+\infty} (u - a)^2 p(u) du = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (u - a)^2 e^{-\frac{(u-a)^2}{2\sigma^2}} du. \quad (*) \end{aligned}$$

Зробимо заміну

$$t = \frac{u-a}{\sigma}.$$

Звідси

$$u - a = t \cdot \sigma, \quad du = \sigma \cdot dt.$$

Продовжимо співвідношення (*):

$$(*) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} dt = -\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t d(e^{-\frac{t^2}{2}}) =$$

$$= -\frac{\sigma^2}{\sqrt{2\pi}} t e^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = (**)$$

Оскільки

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1,$$

то

$$(**) = \sigma^2$$

Отже

$$DX = \sigma^2$$

Виявилось, що параметр σ^2 в нормальному розподілі $N(a, \sigma^2)$ – це дисперсія випадкової величини. Нагадаємо, що a – це математичне сподівання:

$$a = EX, \quad DX = \sigma^2$$

НЕЗАЛЕЖНІ ВИПАДКОВІ ВЕЛИЧИНИ. ДИСПЕРСІЯ НЕЗАЛЕЖНИХ ВИПАДКОВИХ ВЕЛИЧИН

Визначення. Випадкові величини $X_1, X_2, \dots, X_n$ називаються незалежними якщо для будь-яких $t_1, t_2, \dots, t_n$ випадкові події

$$P\left(\bigcap_{i=1}^n (X_i < t_i)\right) = \prod_{i=1}^n P(X_i < t_i).$$

З визначення випливає, що будь-які події, які пов'язані з випадковими величинами $X_1, X_2, \dots, X_n$ будуть незалежними.

Математичне сподівання добутку незалежних випадкових величин

Теорема. Якщо X, Y незалежні випадкові величини. Тоді

$$E(X \cdot Y) = EX \cdot EY$$

Доведення (у випадку дискретної випадкової величини). Нехай маємо розподіл для випадкової величини X :

X	x_1	x_2	$\dots$	x_n	$\dots$
$\mathbb{P}$	p_1^x	p_2^x	$\dots$	p_n^x	$\dots$

В свою чергу, нехай розподіл випадкової величини Y має вигляд:

Y	y_1	y_2	$\dots$	y_k	$\dots$
$\mathbb{P}$	p_1^y	p_2^y	$\dots$	p_k^y	$\dots$

Обчислимо $E(X \cdot Y)$:

$$\begin{aligned}
 E(X \cdot Y) &= \sum_n \sum_k (x_n \cdot y_k) P(X = x_n, Y = y_k) = \\
 &= \sum_n \sum_k (x_n \cdot y_k) \cdot P(X = x_n) \cdot P(Y = y_k) = \\
 &= \sum_n \sum_k (x_n \cdot y_k) \cdot p_n^x \cdot p_k^y = \\
 &= \sum_n x_n \cdot p_n^x \sum_k y_k \cdot p_k^y = EX \cdot EY
 \end{aligned}$$

Дисперсія суми незалежних випадкових величин

Теорема. Якщо X, Y - незалежні випадкові величини. Тоді

$$D(X + Y) = DX + DY$$

Доведення. Позначимо

$$a = EX, \quad b = EY$$

За визначенням дисперсії

$$\begin{aligned} D(X + Y) &= E((X + Y - E(X + Y))^2) = \\ &= E((X + Y - (a + b))^2) = E((X - a) + (Y - b))^2 = \\ &= E((X - a)^2 + 2(X - a)(Y - b) + (Y - b)^2) = \\ &= \underbrace{E(X - a)^2}_{DX} + 2 \underbrace{E(X - a)(Y - b)}_{E(X - a) \cdot E(Y - b)} + \underbrace{E(Y - b)^2}_{DY} = (*) \end{aligned}$$

Оскільки за теоремою про математичне сподівання добутку незалежних випадкових величин

$$E(X - a)(Y - b) = E(X - a)E(Y - b),$$

В той же час

$$E(X - a) = 0, \quad E(Y - b) = 0$$

Тому

$$(*) = DX + DY$$

Приклад 7. Нехай X – біноміальна випадкова величина:

$$P(X = k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}, \quad k=0,1,2,\dots,n$$

Нагадаємо, що розглядається схема Бернуллі, тобто n незалежних експериментів, у кожного з яких може статися «успіх» («У») з ймовірністю p і «неуспіх» («Н») з ймовірністю $q = 1-p$. Випадкова величина X дорівнює кількості «успіхів» («У») в послідовності n незалежних випробувань

Представимо випадкову величину X як суму індикаторів (i – номер випробувань)

$$Y_i = \begin{cases} 1, & \text{якщо здійснився "У"}; \\ 0, & \text{якщо не здійснився "У"} \end{cases}$$

Тоді випадкову величину X можна представити у вигляді

$$X = Y_1 + Y_2 + \dots + Y_n$$

Відповідно до теореми про дисперсію суми незалежних випадкових величин маємо

$$DX = DY_1 + DY_2 + \dots + DY_n = pq + pq + \dots + pq = npq$$

Відповідно до теореми про дисперсію суми незалежних випадкових величин маємо

$$DX = DY_1 + DY_2 + \dots + DY_n = pq + pq + \dots + pq = npq$$

ТЕМА 10. РОЛЬ ГРАНИЧНИХ ТЕОРЕМ В АНАЛІЗІ ВЕЛИКИХ МАСИВІВ ДАНИХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 10

В подальшому, для доведення закону великих чисел нам знадобиться наступні нерівності Маркова та нерівність Колмогорова.

Нерівність Маркова. Нехай $0 \leq X$ - невід'ємна випадкова величина.

Тоді для будь-якого $\varepsilon > 0$ має місце нерівність

$$P(X \geq \varepsilon) \leq \frac{EX}{\varepsilon}$$

Доведення (у випадку дискретної випадкової величини). Нехай маємо розподіл для випадкової величини X :

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

Позначимо через k такий номер, що виконується нерівність

$$x_1 \leq x_2 \leq \dots \leq x_{k-1} < \varepsilon \leq x_k \leq \dots \leq x_n \leq \dots.$$

Тоді

$$EX = \sum_i x_i p_i = \sum_{i=1}^{k-1} x_i p_i + \sum_{i=k} x_i p_i \geq$$

$$\geq \sum_{i=k} x_i p_i \geq \varepsilon \sum_{i=k} p_i = \varepsilon P(X \geq \varepsilon).$$

Нерівність Колмогорова. Нехай X - випадкова величина.

Тоді для будь-якого $\varepsilon > 0$ має місце нерівність

$$P(|X - a| \geq \varepsilon) \leq \frac{DX}{\varepsilon^2}$$

Доведення Оскільки

$$0 \leq Y = (X - a)^2$$

-невід'ємна випадкова величина, де $a = EX$, то за нерівністю Маркова для будь-якого $\varepsilon > 0$ маємо співвідношення

$$\begin{aligned} P(|X - a| \geq \varepsilon) &= P((X - a)^2 \geq \varepsilon^2) = \\ &= P(Y \geq \varepsilon^2) \leq \frac{EY}{\varepsilon^2} = \frac{DX}{\varepsilon^2} \end{aligned}$$

ЗАКОН ВЕЛИКИХ ЧИСЕЛ

Збіжність послідовності випадкових величин характеризує їх наближення в деякому розумінні до деякої випадкової величини (зокрема, можливо і до константи. Будемо розглядати збіжність за ймовірністю та з ймовірністю одиниця (збіжність майже напевно).

Збіжність за ймовірністю

(Convergence in probability)

Визначення. Кажуть, що послідовність випадкових величини $Y_1, Y_2, \dots, Y_n$ збігається за ймовірністю до випадкової величини Y , якщо для будь-якого $\varepsilon > 0$

$$P(|Y_n - Y| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0$$

і позначається

$$Y_n \xrightarrow{P} Y, \quad n \rightarrow \infty.$$

Закон великих чисел Чебишева

(The Law of Large Numbers)

Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність незалежних випадкових величин, дисперсії яких обмежені деякою константою C :

$$DX_n \leq C.$$

Тоді

$$\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n EX_i \xrightarrow{P} 0, \quad n \rightarrow \infty. \quad (*)$$

Доведення. Позначимо

$$Y_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad n = 0, 1, 2, \dots$$

Оскільки математичне сподівання суми випадкових величин дорівнює сумі математичних сподівань цих величин, то

$$\frac{1}{n} \sum_{i=1}^n EX_i = EY_n$$

В нових позначеннях (*) можна переписати у вигляді

$$Y_n - EY_n \xrightarrow{P} 0, \quad n \rightarrow \infty.$$

Відповідно до нерівності Чебишева маємо

$$P(|Y_n - EY_n| \geq \varepsilon) \leq \frac{DY_n}{\varepsilon^2} \leq (**)$$

Для незалежних випадкових величин дисперсія суми цих величин дорівнює сумі їх дисперсій

$$DY_n = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) =$$

$$= \frac{1}{n^2} \left(\sum_{i=1}^n DX_i\right)$$

Продовжуючи (**) і зважаючи на обмеженість дисперсій $X_1, X_2, \dots, X_n, \dots$ отримуємо нерівність

$$\frac{1}{\varepsilon^2 n^2} \left(\sum_{i=1}^n DX_i\right) \leq \frac{1}{\varepsilon^2 n^2} \underbrace{(C + C + \dots + C)}_n = \frac{C}{\varepsilon^2 n}.$$

Очевидно, що

$$\frac{C}{\varepsilon^2 n} \xrightarrow{n \rightarrow \infty} 0,$$

Звідки отримуємо збіжність (*). Теорема Чебишева доведена.

З теореми Чебишева випливає закон величин чисел для однаково розподілених випадкових величин. Випадкові будуть однаково розподіленими, якщо вони мають однакові функції розподілу. Зокрема, для дискретних випадкових величин однаково розподілені випадкові величини мають однакові ряди розподілу. Прикладом однаково розподілених випадкових величин

можуть слугувати величини в представлені біноміальної випадкової величини індикаторами «Успіхів» в n незалежних випробувань (схема Бернуллі).

Нагадаємо, що розглядається схема Бернуллі, тобто n незалежних експериментів, у кожного з яких може статися «Успіх» з ймовірністю p і «Неуспіх» з ймовірністю $q = 1-p$. Випадкова величина X дорівнює кількості «Успіхів» в послідовності n незалежних випробувань.

Закон великих чисел
для однаково розподілених випадкових величин
(The Law of Large Numbers
for Identically Distributed Random Variables)

Нехай $X_1, X_2, \dots, X_n$ незалежні однаково розподілені випадкові величини з математичним сподіванням

$$a \equiv EX_1 = EX_2 = \dots = EX_n = \dots.$$

Тоді

$$\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{P} a, \quad n \rightarrow \infty.$$

Розглянемо важливий окремий випадок цієї теореми для схеми Бернуллі. Нагадаємо, що розглядаються n незалежних експериментів, у кожного з яких може статися «Успіх» з ймовірністю p і «Неуспіх» з ймовірністю $q = 1-p$. Випадкова величина X дорівнює кількості «Успіхів» в послідовності n незалежних випробувань. Кількість «Успіхів» в цій послідовності незалежних

випробувань – це біноміальна випадкова величину X . Її можна представити, як суму

$$X = Y_1 + Y_2 + \dots + Y_n$$

незалежних індикаторів (i – номер випробувань)

$$Y_i = \begin{cases} 1, & \text{якщо здійснився "Успіх";} \\ 0, & \text{якщо не здійснився "Успіх"} \end{cases}$$

Тоді випадкова величина

$$v_n = \frac{Y_1 + Y_2 + \dots + Y_n}{n}$$

являє собою відносну частоту появи «Успіхів» в n незалежних випробувань.

Теорема Бернуллі. Для будь-якого $\varepsilon > 0$

$$P(|v_n - p| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0$$

Смисл цього твердження наступний. Якщо під «Успіхом» розуміти, що відбувається деяка подія A , то $p = P(A)$ – це ймовірність цієї події. Таким чином **введене поняття ймовірності відповідає нашому інтуїтивному уявленню про ймовірність як границі частоти**

$$v_n \approx P(A).$$

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 11
Приклад комп'ютерного моделювання (computer simulation)

Метод Монте-Карло

Припустимо, що необхідно обчислити інтеграл

$$\int_0^1 \varphi(u) du.$$

Розглянемо послідовність $X_1, X_2, \dots, X_n \dots$ незалежних рівномірно на відрітку $[0, 1]$ розподілених випадкових величин з щільністю

$$P(u) = \begin{cases} 1, & \text{якщо } u \in [0, 1] \\ 0, & \text{якщо } u \notin [0, 1] \end{cases}$$

Тоді

$$E\varphi(X_n) = \int_{-\infty}^{+\infty} \varphi(u)p(u)du = \int_0^1 \varphi(u) du$$

Згідно закону великих чисел для однаково розподілених випадкових величин

$$\frac{\varphi(X_1) + \varphi(X_2) + \dots + \varphi(X_n)}{n} \xrightarrow{P} \int_0^1 \varphi(u) du, \quad n \rightarrow \infty. \quad (*)$$

Сформулюємо відповідний алгоритм обчислення інтеграла методом Монте-Карло.

Сформулюємо відповідний алгоритм обчислення інтеграла методом Монте-Карло.

Крок 1. Генеруємо послідовність чисел

$$x_1, x_2, \dots, x_n,$$

які є реалізацією послідовності незалежних випадкових величин, що мають рівномірне на відріжку $[0, 1]$ розподіл (це можна зробити, наприклад, за допомогою Excel).

Крок 2. З урахуванням (*) вважаємо

$$\int_0^1 \varphi(u) du \approx \frac{\varphi(x_1) + \varphi(x_2) + \dots + \varphi(x_n)}{n}$$

Зауважимо, що метод Монте-Карло є особливо ефективним при обчисленні інтегралів великої кратності. Для таких інтегралів методи прямих обчислень можуть бути непридатними. Аналогічно можна застосовувати метод Монте-Карло для обчислення моделей економічних систем.

Збіжність з ймовірністю одиниця (Almost surely convergence)

(майже скрізь = almost surely = майже напевно)

Збіжність за ймовірністю не дозволяє відстежувати поведінку реалізацій послідовності випадкових величин. Можна оцінити тільки близькість однієї випадкової події до іншої. Зробити висновок про долю (ймовірність) траєкторій, що мають визначені властивості – проблематично. Збіжність з ймовірністю одиниця дозволяє зробити висновки про траєкторії послідовності випадкових подій. Важливо зазначити, що **якщо послідовність випадкових величин збігається з ймовірністю одиниця, то вона збігається за ймовірністю.**

Визначення. Кажуть, що послідовність випадкових величини $Y_1, Y_2, \dots, Y_n$ збігається з імовірністю одиниця (майже завжди) до випадкової величини Y , якщо

$$P \left(Y_n \xrightarrow[n \rightarrow \infty]{} Y \right) = 1$$

Це означає, що всі («майже всі») траєкторії

$$Y_1(\omega), Y_2(\omega), \dots, Y_n(\omega), \dots$$

збігаються до $Y(\omega)$.

Посилений закон великих чисел для однаково розподілених випадкових величин (*Strong Law of Large Numbers for Independent Identically Distributed Random Variables*)

Нехай $X_1, X_2, \dots, X_n$ незалежні однаково розподілені випадкові величини з математичним сподіванням

$$a \equiv EX_1 = EX_2 = \dots = EX_n = \dots.$$

Тоді

$$P \left(\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{} a \right) = 1.$$

Приклад прийняття рішення про участь у проєкті.

Припустимо, що проєкт здійсниться дорівнює 0,7, а що проєкт зазнає невдачі дорівнює 0,3. Витрати на реалізацію проєкта становлять 100'000 грошових одиниць. Запланований валовий дохід дорівнює 130'000 (тобто чистий прибуток - 30'000). Таким чином, в разі здійснення проєкту маємо фінансовий результат в +30'000, в іншому випадку фінансовий результат від'ємний: -100'000.



Приклад прийняття рішення про участь у проєкті (рішення).

Розглянемо випадкову величину X , яка дорівнює фінансовому результату при здійсненні проєкту:

$$X = \begin{cases} +30'000 & \text{з ймовірністю } 0,7 \text{ якщо здійсниться проєкт;} \\ -100'000 & \text{з ймовірністю } 0,3 \text{ якщо не здійснився проєкт} \end{cases}$$

Уявимо, що ситуація прийняття рішень про участь у проєкті повторюється n

разів. Послідовність фінансових результатів $X_1, X_2, \dots, X_n$ являє собою послідовність незалежних однаково розподілених випадкових величин, які мають розподіл, що співпадає з розподілом X . Неважко підрахувати, що їх (однакове) математичне сподівання дорівнює

$$a = (+30'000) \cdot 0,7 + (-100'000) \cdot 0,3 = -9000.$$

Згідно посиленому закону великих чисел

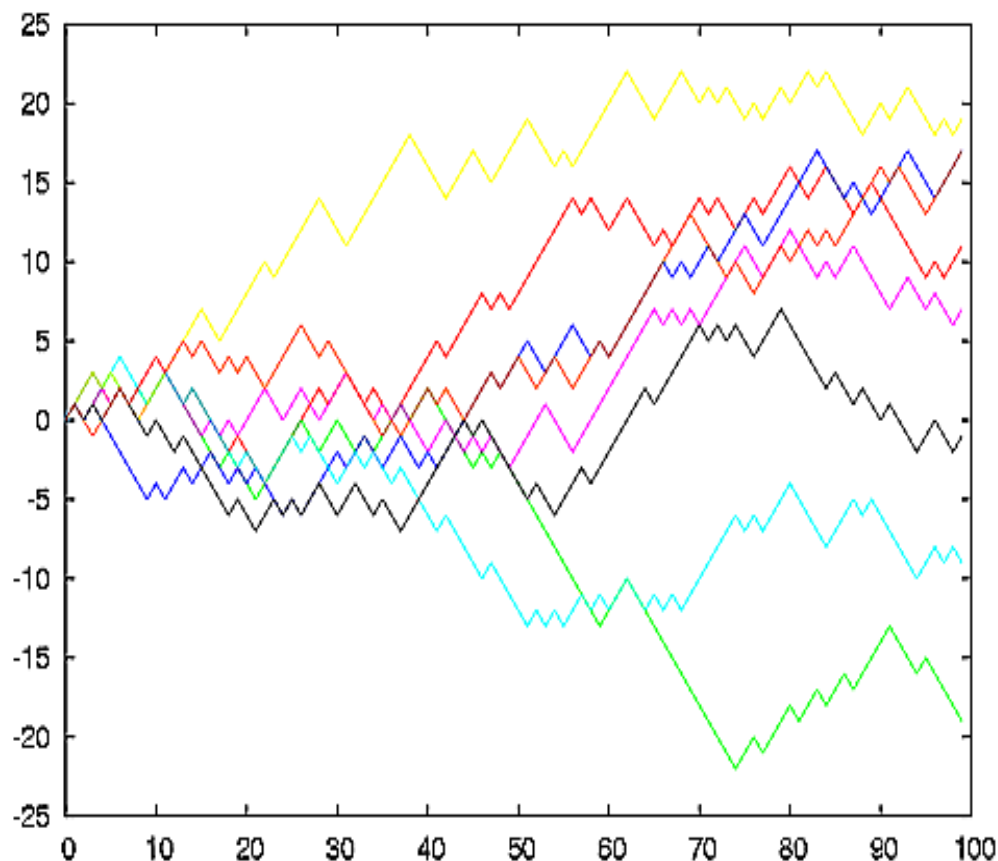
$$P \left(\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{n \rightarrow \infty} -9'000 \right) = 1.$$

Тобто в середньому будемо мати 9'000 збитків. Тому природньо відмовитися від реалізації проєкту.

Випадкові блукання. Модель розорення (банкрутства)

Теорія випадкових блукань використовується в економіці та фінансах для опису непередбачуваних змін ринкових цін. Згідно з цією теорією, ринкові ціни, як і доходи від інвестицій, змінюються випадковим чином, без чітко прогнозованого напрямку. Це пов'язане з гіпотезою ефективного ринку (perfect market), яка стверджує, що всю доступну інформацію вже враховано в поточній ціні активів, а отже, майбутні зміни є результатом нових випадкових подій.

Ідея випадкових блукань у фінансах та маркетингу показує, що ринок не можна стабільно передбачити, а інвесторам краще розглядати його як систему, де успіх здебільшого залежить від довгострокової стратегії та диверсифікації, а не від спроб прогнозувати короткострокові коливання.



Візуалізація декількох випадкових блукання, які виходять з нуля ($S_0 = 0$). По осі абсцис – число підкидань, по осі ординат – сума коштів.

Mosteller F. Fifty Challenging Problems in Probability with Solutions. -1987. Dover Publications. 96 p.

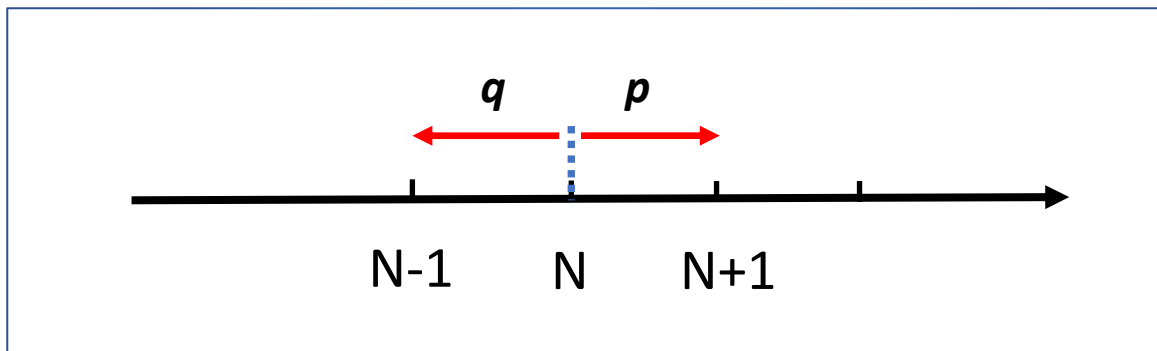
Нехай $X_1, X_2, \dots, X_n$ - послідовність незалежних однаково розподілених випадкові величини:

$$X_n = \begin{cases} 1, & \text{з ймовірністю } p \\ -1, & \text{з ймовірністю } q = 1 - p \end{cases}$$

Позначимо їх суми

$$S_n = X_1 + X_2 + \dots + X_n, \quad n = 0, 1, 2, \dots$$

Послідовність $\{S_n, n = 0, 1, 2, \dots\}$ можна розглядати, як найпростіше блукання частинки на прямій з кроком в одиницю.



Подивимось на випадкове блукання з іншої точки зору. Уявимо, що граємо в таку гру:

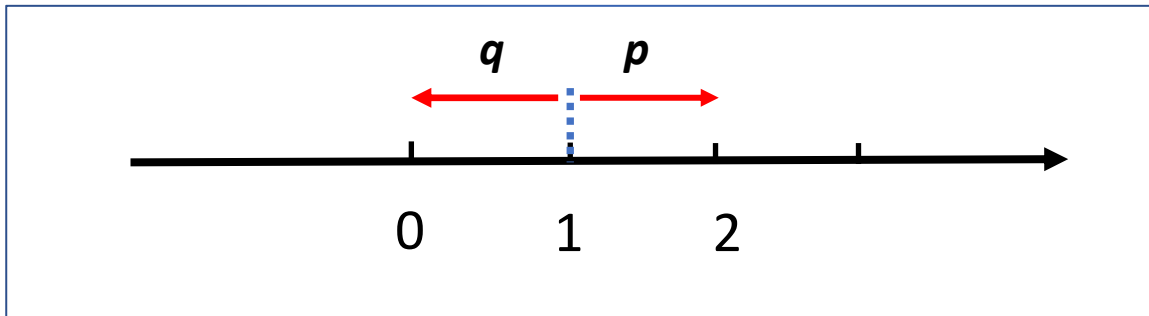
підкидуємо монету; якщо випадає «Герб», то гравець №1 отримує винагороду в «+1» одиницю. Якщо випадає «Решка», то виграв в одну одиницю отримує гравець №2 (відповідно гравець №1 отримує програш у розмірі «-1»). Припустимо, що на початку перший гравець має N грошових одиниць, а другий – M . Постає питання: чи буде відрізнятися ймовірність виграву першого гравця від другого? І, якщо буде, то наскільки?

З точки зору опису гри, яка наведена вище, S_n – є сума грошових коштів після n -ого підкидання, а S_0 – початкова сума грошей першого гравця. Перший момент, коли S_n дорівнює нулю - це момент розорення гравця №1.

Банкрутство у грі проти гравця з необмеженими грошовими ресурсами

Розглянемо гру, в якій гравець, що має N грошових одиниць ($S_0 = N$), грає проти гравця з необмеженими грошовими ресурсами ($M = +\infty$).

Почнемо з обчислення ймовірності банкрутства гравця №1, якщо він має одну грошову одиницю. Позначимо цю ймовірність через $P(1) \equiv x$. Послідовність коштів, яку гравець має після кожної гри можемо розглядати як рух частинки на прямій: кожний раз з ймовірністю p вона зміщується вправо, або з ймовірністю $1 - p = q$ вліво.



В перший момент частинка зміститься вліво з ймовірністю p і гравець розориться, або зміститься вправо з ймовірністю $1 - p = q$ і буде мати 2 одиниці ($X_1 = 1; S_1 = 2$). Ймовірність того, що гравець з 2 одиницями розориться позначимо $P(2)$. Очевидно, що

$$P(2) = P\{\text{настане такий момент } k, \text{ що } X_2 + X_3 + \dots + X_k = -2\}$$

Оскільки X_1 не залежить від $X_2 + X_3 + \dots + X_k$, то

$$\begin{aligned} P(1) &= P(X_1 = -1) + \\ &+ P\{X_1 = 1 \text{ та настане такий момент } k, \text{ що } X_2 + X_3 + \dots + X_k = -2\} = \\ &= 1 - p + p \cdot P(2) \end{aligned}$$

Маємо

$$P(1) = 1 - p + p \cdot P(2) \quad (*)$$

Якщо частинка знаходиться в точці 2, то для того щоб відбулося розорення гравця треба частинці спочатку попасти в точку 1, а потім в точку 0. Тобто, гравець, який має 2 одиниці, розориться у випадку, якщо одночасно стануться дві події

$$A = \{\text{настане такий момент } k, \text{ що } X_2 + X_3 + \dots + X_k = -1\}$$

$$B = \{\text{настане такий момент } i > k, \text{ що } X_{k+1} + X_{k+2} + \dots + X_i = -1\}$$

Не складно зрозуміти, що

$$P(A) = P(B) = x$$

Через те, що випадкові величини $X_2 + X_3 + \dots + X_k$ та $X_{k+1} + X_{k+2} + \dots + X_i$ незалежні

$$P(2) = P(A \cap B) = P(A) \cdot P(B) = x^2 \quad (**)$$

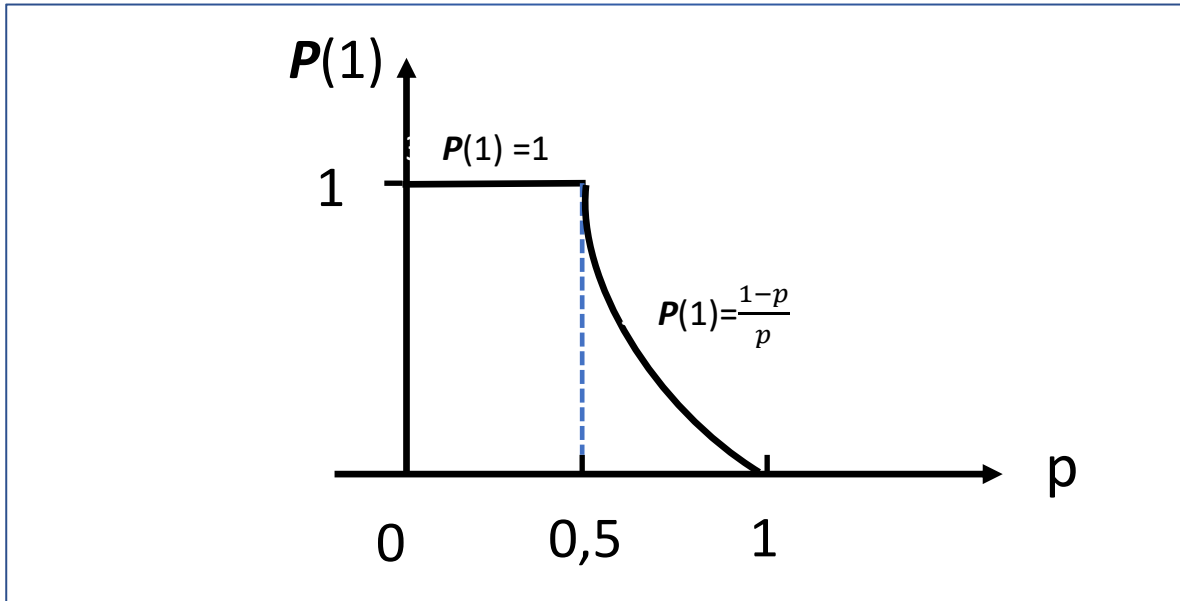
Тоді співвідношення (*) призводить до квадратного рівняння

$$x = 1 - p + p \cdot x^2$$

Якщо $p = 0$, то маємо розв'язок $x_1 = 1$. У випадку $p > 0$ неважко знайти розв'язки цього рівняння:

$$x_1 = 1, \quad x_2 = \frac{1}{p} - 1$$

При $p < \frac{1}{2}$ другий розв'язок $x_2 > 1$ не задовольняє умовам задачі (x — це ймовірність). При $p = \frac{1}{2}$ розв'язки співпадають. Отже, при $0 \leq p \leq \frac{1}{2}$ маємо один розв'язок $x_1 = 1 = P(1)$. Це означає, що при $0 \leq p \leq \frac{1}{2}$ гравець №1 збанкрутує з ймовірністю одиниця, при $\frac{1}{2} < p \leq 1$ гравець №1 збанкрутує з ймовірністю $P(1) = \frac{1}{p} - 1 = \frac{1-p}{p} = \frac{q}{p}$.



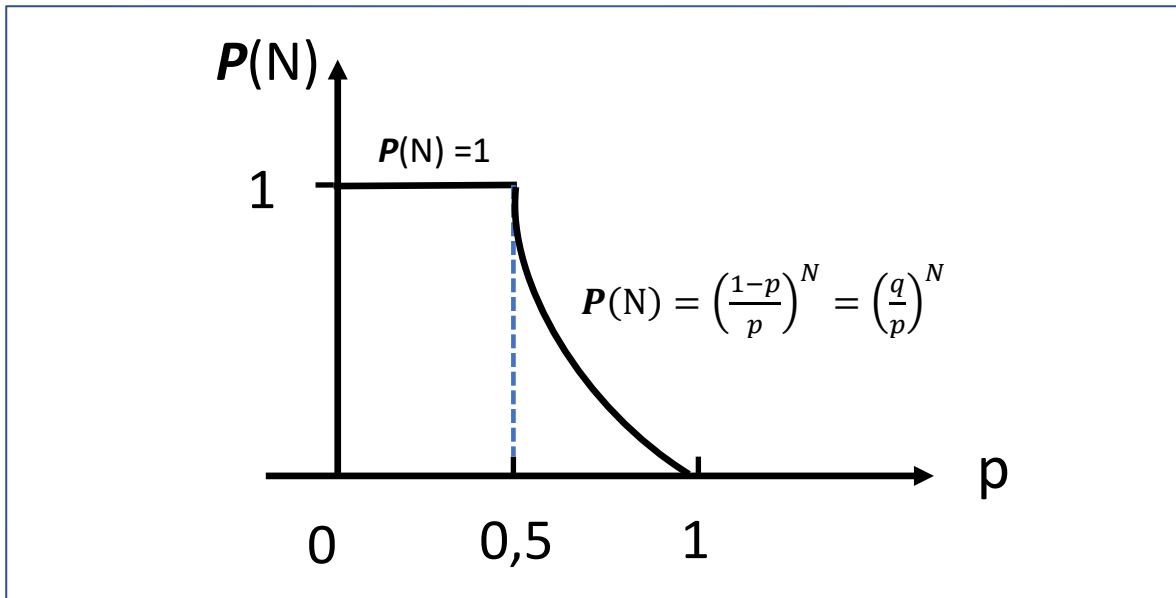
Позначимо ймовірність того, що гравець з N одиницями розориться позначимо $P(N)$. Якщо повторити міркування які призводять до співвідношення (**), тоді отримаємо

$$P(N) = P(N-1)P(1)$$

Звідси

$$P(N) = P(N-1)P(1) = P(N-2)P(1)^2 = \dots = P(1)^N = \left(\frac{q}{p}\right)^N.$$

Тоді графік залежності банкрутства гравця №1 в разі, якщо він володіє на початку N грошових одиниць має наступний вид

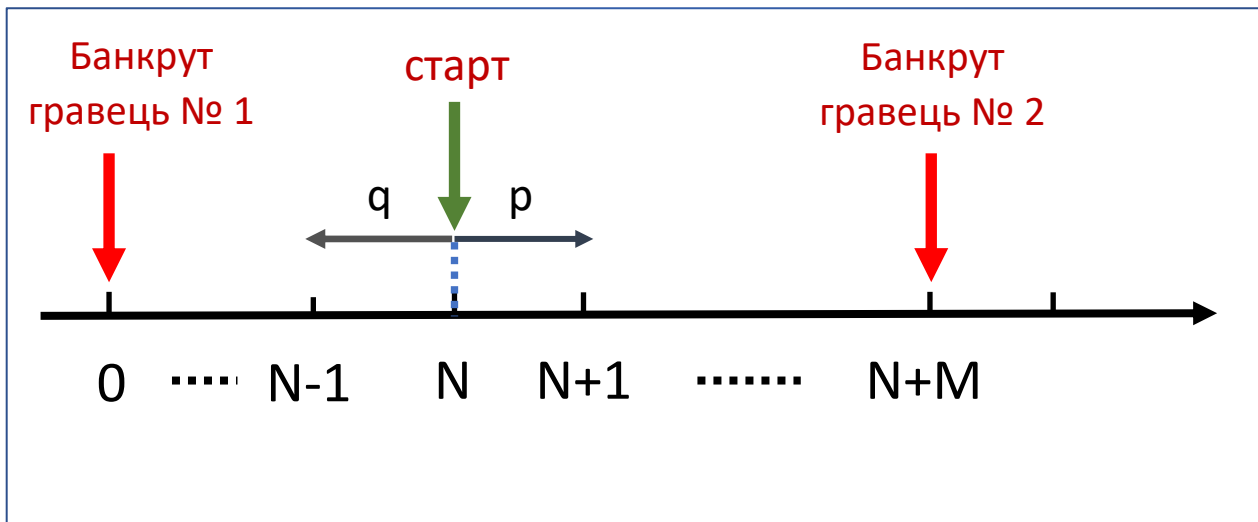


Звертає на себе увагу наступний парадокс. А саме, якщо маємо «нейтральну» гру з $p = \frac{1}{2}$, коли ймовірність виграшу та програшу однакові, то гравець №1 збанкрутує з ймовірністю одиниця. При цьому згідно посиленому закону великих чисел середній виграш (програш) буде прямувати до нуля (оскільки $EX_1 = EX_2 = \dots = EX_n = \dots = 0$):

$$P \left(\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{n \rightarrow \infty} 0 \right) = 1$$

Банкрутство у грі гравців з обмеженими грошовими ресурсами

Нехай гравець №1 має N грошових одиниць, а гравець №2 – M грошових одиниць. Гра триває до моменту банкрутства одного з гравців. Яка ймовірність виграшу в цієї гра першого гравця (відповідно програшу другого)?



«0» - стан банкрутства гравця № 1; «**M+N**» – стан банкрутства гравця № 2

Вище було доведено, що ймовірність програшу гравця №1 в грі проти гравця №2 з необмеженими ресурсами дорівнює

$$\left(\frac{q}{p}\right)^N$$

Тобто це доля тих траєкторій, що досягають нуля. Цю множину траєкторій розіб'ємо на такі, що доходять до точки **M+N** (точки банкрутства гравця №2) і такі, що не доходять до цієї точки. Перша підмножина траєкторій описують ситуацію виграшу гравця №1, а друга підмножина – його програшу. Нехай **L** – ймовірність програшу першого гравця другому, яка дорівнює долі траєкторій, що не доходять до точки **M+N** (точки банкрутства гравця №2. Тоді

$$\left(\frac{q}{p}\right)^N = L + (1 - L) \cdot \left(\frac{q}{p}\right)^{N+M}$$

Якщо **W** = 1 – **L** означає ймовірність виграшу першого гравця, то

$$W = \frac{1 - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^{N+M}} \quad (**) \quad$$

В окремому випадку $p = q = \frac{1}{2}$ (розкриваючи невизначеність 0/0 за правилом Лопітала) маємо

$$W = \frac{N}{N + M}$$

Наприклад, якщо в «нейтральну» грі ($p = q = \frac{1}{2}$) гравець №1 має в два рази більше коштів ($N = 2M$) ніж гравець №2, то ймовірність його виграшу буде $\frac{2}{3}$. Відповідно ймовірність виграшу гравця №2 становитиме $\frac{1}{3}$. Отже ймовірність виграшу першого гравця буде в два рази більше ймовірності виграшу другого гравця.

Задача про «обережного» та «ризикового» гравця

У чоловіка, що знаходиться в Лас-Вегасі залишилося 30 доларів. Білет до дому коштує 60 доларів. Він розмірковує над двома стратегіями поведінки в казино при грі в рулетку. Перша – стратегія «обережного» гравця: робити ставку на «парне» (або «непарне») по одному долару. Друга – стратегія «ризикового» гравця: зробити відразу ставку в 30 доларів і, таким чином, закінчити гру одразу ж. Питання: дотримуючись якої стратегії ймовірність уїхати додому буде більше ?

Розв'язання. Нехай застосовується стратегія «ризикового» гравця. Оскільки «парних» чисел в рулетці 18, а всього чисел 38 (є ще дві позиції «зеро», ZERO), ймовірність купити білет додому буде дорівнювати $\frac{18}{38} \approx 0,474$.

Обчислимо ймовірність виграшу (тобто ймовірність покупки білету додому), якщо застосовувати стратегію «обережного гравця».

В цьому випадку маємо ситуацію задачу про банкрутство у грі гравців з обмеженими грошовими ресурсами з

$$N = 30, \quad M = 30, \quad p = \frac{18}{38}, \quad q = \frac{20}{38}$$

За формулою (**) отримуємо ймовірність виграшу першого гравця, тобто ймовірність купівлі білету додому

$$W = \frac{1 - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^{N+M}} = \frac{1 - \left(\frac{20}{18}\right)^{30}}{1 - \left(\frac{20}{18}\right)^{60}} = \frac{23,59 - 1}{556,48 - 1} \approx 0,041$$

Отже, шансів поїхати додому у «обережного» гравця майже в 11 разів більше ніж у «ризикового» гравця.

Закон арксинуса. Нехай $\{S_n, n = 0, 1, 2, \dots\}$ - найпростіше симетричне (тобто $p = q = \frac{1}{2}$) блукання з шагом в одиницю, а τ_n - доля часу перебування частинки на додатній стороні (тобто кількість моментів i , для яких $S_i > 0$). Тоді твердження закону арксинусу має наступний вигляд.

Теорема Ймовірність того, що доля часу, який частинка проводить на додатній стороні, менше t , прямує до $\frac{2}{\pi} \cdot \arcsin \sqrt{t}$, тобто для $0 \leq t \leq 1$

$$P(\tau_n < t) \xrightarrow{n \rightarrow \infty} \frac{2}{\pi} \cdot \arcsin \sqrt{t}. \quad (*)$$

Наприклад, скориставшись таблицями арксинусу і тими обставинами, що швидкість збіжності в (*) достатньо велика, отримаємо

$$P(\tau_n < 0,024) \approx 0,1.$$

Джерело: Székely G.J. Paradoxes in Probability Theory and Mathematical Statistics. 1986. Reidel, Kluwer Academic Publishers. 264 p.

Таким чином, зокрема, якщо $n = 1000$, то приблизно одна з десяти частинок буде перебувати на додатній стороні не більш ніж 24 одиниці часу, а на від'ємній стороні – не менш ніж 976 одиниць часу. Оскільки випадкове блукання симетричне, то і на від'ємній стороні приблизно одна з десяти частинок буде перебувати не більш ніж 24 одиниць часу, а на додатній стороні не менш ніж 975 одиниць часу. Тобто приблизно кожна з п'ята частка майже весь час (97,6%) буде перебувати тільки з однієї сторони (додатній або від'ємній).

ГРАНИЧНІ ТЕОРЕМИ ТЕОРІЇ ЙМОВІРНОСТЕЙ

Центральна гранична теорема (Central Limit Theorem)

Центральна гранична теорема – це будь-яка теорема, що ставить умови, при яких функція розподілу суми індивідуально малих випадкових величин з ростом числа доданків сходиться до нормальної функції розподілу.

Центральна гранична теорема підтверджує наступне: якщо результат випадкового експерименту визначається великим числом випадкових факторів, вплив кожного з яких дуже малий, то такий експеримент добре апроксимується нормальним розподілом $N(a, \sigma^2)$ з параметрами a, σ^2 , підібраними відповідним чином.

Сформулюємо центральну граничну теорему в загальному випадку необов'язково однаково розподілених випадкових величин.

Позначення. Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність незалежні випадкові величини;

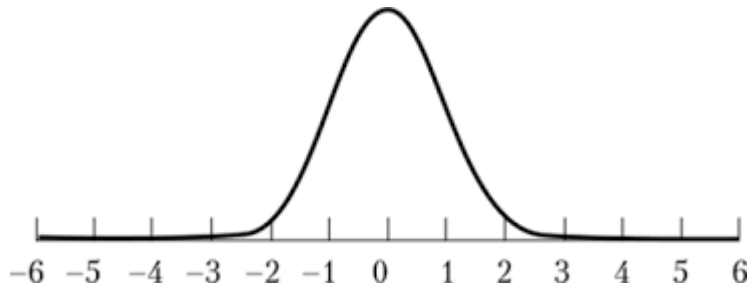
$$a_k = EX_k, \quad \sigma_k^2 = DX_k, \quad T_k^3 = E |X_k - a_k|^3,$$

Позначимо функцію розподілу центрованих та нормованих сум цих випадкових величин

$$F_n(t) = P \left\{ \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n EX_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} < t \right\}$$

Функцію розподілу стандартної нормальної випадкової величини (тобто з нульовим математичним сподіванням та одиничною дисперсією $N(0,1)$) як

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{u^2}{2}} du$$



Щільність розподілу стандартного нормального розподілу

Умова Ляпунова:

$$\lim_{n \rightarrow \infty} \frac{\sqrt[3]{T_1^3 + T_2^3 + \dots + T_n^3}}{\sqrt[2]{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}} = 0$$

Твердження (Центральна гранична теорема)

Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність незалежні випадкові величини. Виконано умова Ляпунова. Тоді для кожного t має місце збіжність

$$F_n(t) \xrightarrow{n \rightarrow \infty} \Phi(t)$$

Сформулюємо центральну граничну теорему у випадку однаково розподілених доданків.

Позначення. Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність незалежні однаково розподілених випадкових величини;

$$a = EX_k, \quad \sigma^2 = DX_k, \quad T^3 = E|X_k - a|^3,$$

Функцію розподілу центрованих та нормованих сум цих випадкових величин позначимо наступним чином

$$F_n(t) = P \left\{ \frac{\sum_{i=1}^n X_i - na}{\sigma \sqrt{n}} < t \right\}$$

Твердження (Центральна гранична теорема для однаково розподілених доданків) . Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність незалежні однаково розподілених випадкових величини. Тоді для кожного t має місце збіжність

$$F_n(t) \xrightarrow{n \rightarrow \infty} \Phi(t)$$

Доведення. Перевіримо умову Ляпунова

$$\lim_{n \rightarrow \infty} \frac{\sqrt[3]{T_1^3 + T_2^3 + \dots + T_n^3}}{\sqrt[2]{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}} = \lim_{n \rightarrow \infty} \frac{\sqrt[3]{nT^3}}{\sqrt[2]{n\sigma^2}} = \frac{T}{\sigma} \lim_{n \rightarrow \infty} \frac{1}{\sqrt[6]{n}} = 0$$

Отже з центральної граничної теореми в загальному випадку впливає центральна гранична теорема у випадку однаково розподілених доданків.

Наслідок. Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність незалежні однаково розподілених випадкових величини. Тоді для будь яких α і β має місце приблизна рівність

$$P \left\{ \alpha < \frac{\sum_{i=1}^n X_i - na}{\sigma \sqrt{n}} < \beta \right\} \approx \Phi(\beta) - \Phi(\alpha)$$

де $\Phi(t)$ — функцію розподілу стандартної нормальної випадкової величини, значення якої можна знайти в таблицях.

В окремому випадку, коли Y_n — кількість «успіхів» в n незалежних випробувань (тобто Y_n — біноміальна випадкова величина, яку, як відомо можна представити, як суму незалежних двозначних випадкових величин

$$Y_n = X_1 + X_2 + \dots + X_n$$

Маємо теорему Муавра – Лапласа.

Теорема Муавра – Лапласа Нехай Y_n – кількість «успіхів» в n незалежних експериментів (схема Бернуллі), у кожного з яких може статися «Успіх» з ймовірністю p і «Неуспіх» з ймовірністю $q = 1-p$. Тоді для будь яких α і β

$$P\left\{\alpha < \frac{Y_n - np}{\sqrt{npq}} < \beta\right\} \xrightarrow{n \rightarrow \infty} \Phi(\beta) - \Phi(\alpha)$$

Наближення стають кращими з ростом n . В практиці наближення вважають непоганими, якщо виконуються нерівності

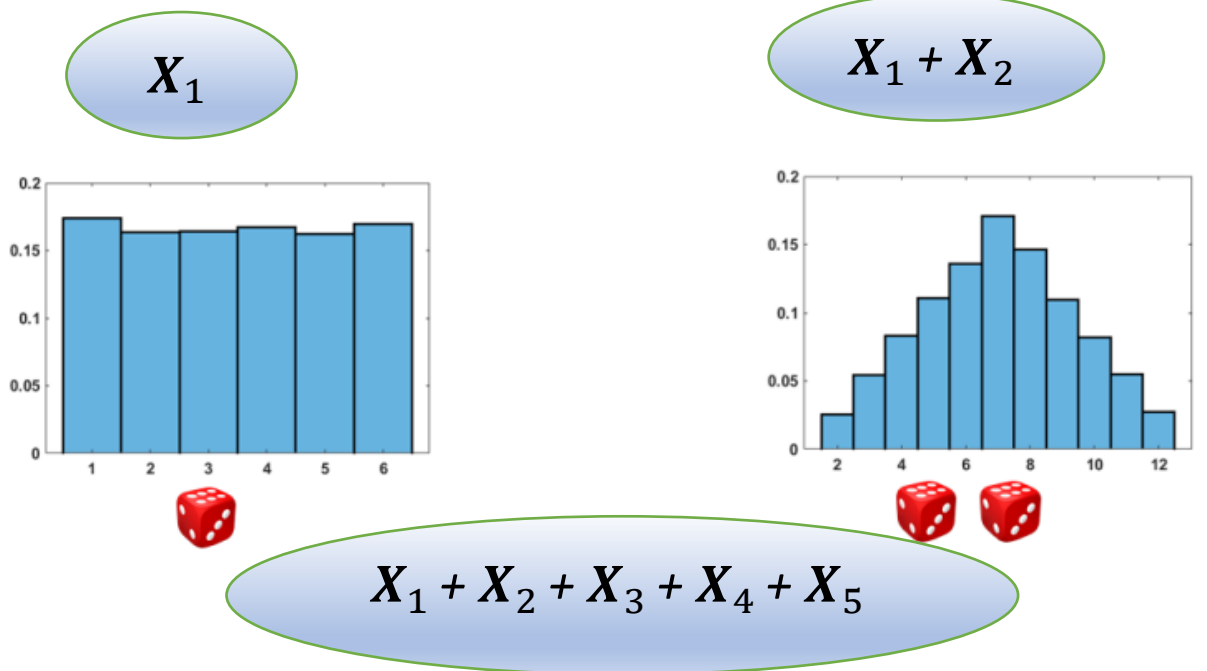
$$np \geq 5, \quad nq \geq 5$$

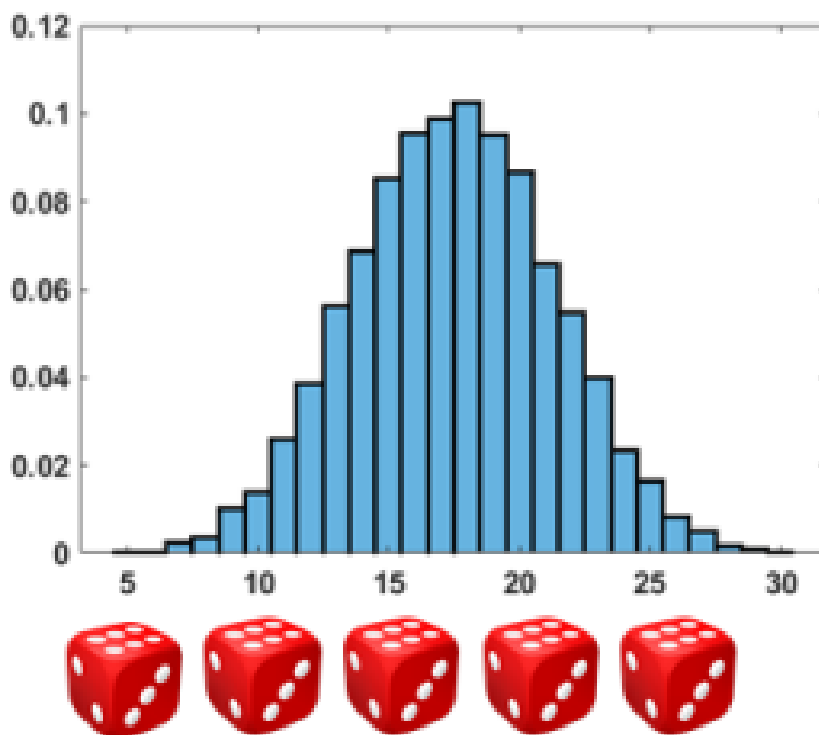
Джерело:

Buldygin V.V., Solntsev S.A. *Asymptotic Behaviour of Linearly Transformed Sums of Random Variables*. 1997. Kluwer Academic Publishers, 500 p.

Buldygin V.V., Solntsev S.A. *The strong law of large numbers for sums of independent random vectors with operator normalizations. Theory of Probability & Its Applications*. 1988. 32 (2), 243-256 pp.

Buldygin V.V., Solntsev S.A. *The contraction principle and the strong law of large numbers for weighted sums. Theory of Probability & Its Applications*. 1987. 31 (3), 452-462 pp.





Джерело: Chan S.H. *Introduction to Probability for Data Science*. 2021. Michigan Publishing Services. 704 p.

Випадкові вектори

Визначення. Сукупність n випадкових величин $(Z_1, Z_2, \dots, Z_n)$, які задані на одному ймовірнісному просторі, називають n -вимірною випадковою величиною, або n -вимірним вектором.

З метою спрощення будемо розглядати двовимірний 2-вимірні вектори.

Визначення. Функцією розподілу випадкового вектору (X, Y) або спільної функцією розподілу випадкових величин (X, Y) називається

$$F(x, y) = P(X < x, Y < y)$$

Властивості функції розподілу:

1. $0 \leq F(x, y) \leq 1$
2. Функція $F(x, y)$ — неспадна по кожному з аргументів x і y .

3. $F(-\infty, y) = F(x, -\infty) = 0$

4. $F(-\infty, +\infty) = 1$

5. Функція $F(x, y)$ – неперервна зліва по кожному з аргументів x і y .

Двовимірною випадковою величиною (X, Y) називають дискретну, якщо X, Y – дискретні випадкові величини. Для неї складається таблиця розподілу – аналог ряду розподілу для одновимірної випадкової величини.

		Y				
X		y_1	y_2	...	y_k	...
	x_1	p_{11}	p_{12}	...	p_{1k}	...
	x_2	p_{21}	p_{22}	...	p_{2k}	...
	...	...	...	...	...	...
	x_n	p_{n1}	p_{n2}	...	p_{nk}	...
	...	...	...	...	...	...

Приклад. Проводяться два постріли в мішень. При кожному пострілі ймовірність попадання p , ймовірність промаху $q = 1 - p$. Випадкова величина X – число влучень при першому пострілі, Y – число влучень при другому пострілі. Знайдемо закон розподілу вектору (X, Y) .

		Y	
X		$y_1 = 0$	$y_2 = 1$
	$x_1 = 0$	q^2	$q \cdot p$
	$x_2 = 1$	$p \cdot q$	p^2

Двовимірною випадковою величиною називають неперервну, якщо X, Y – безперервні випадкові величини і її функцію розподілу можна представити у вигляді

$$F(x, y) = P(X < x, Y < y) = \int_{-\infty}^x \int_{-\infty}^y p(t, u) dt du$$

Функція $p(t, u)$ називається щільністю розподілу випадкового вектора (X, Y) .

Властивості щільності розподілу:

1. $p(t, u) \geq 0$

2. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(t, u) dt du = 1.$

3. $p(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$

Випадкові вектори з незалежними компонентами

Дискретні вектори. Нехай X і Y незалежні та дискретні з рядами розподілу

X	x_1	x_2	...	x_n	...
P	p_1^X	p_2^X	...	p_n^X	...

та

Y	y_1	y_2	...	y_k	...
P	p_1^Y	p_2^Y	...	p_k^Y	...

Тоді для випадкового вектору (X, Y) маємо

$$p_{nk} = P(X = x_n, Y = y_k) = P(X = x_n) \cdot P(Y = y_k) = p_n^X \cdot p_k^Y$$

Неперервні вектори. Нехай X і Y незалежні випадкові величини з функціями розподілу

$$F_X(t), \quad F_Y(u),$$

мають неперервний розподіл; відповідні щільності

$$p^X(t), \quad p^Y(u);$$

вектор (X, Y) з функцією розподілу $F(x, y)$ має неперервний розподіл з щільністю

$$p(x, y).$$

Тоді

$$p(x, y) = p^X(x) \cdot p^Y(y)$$

Дійсно (властивість 3 багатовимірної щільності),

$$p(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial^2 (F_X(x) \cdot F_Y(y))}{\partial x \partial y} = \frac{\partial F_X(x)}{\partial x} \cdot \frac{\partial F_Y(y)}{\partial y} = p^X(x) \cdot p^Y(y).$$

Коваріація

Визначення. Коваріацією випадкових величин X і Y називається

$$\text{Cov}(X, Y) \stackrel{\text{def}}{=} E(X - EX)(Y - EY)$$

Властивості коваріації

$$\boxed{1.} \quad \text{Cov}(X, Y) = EXY - EX \cdot EY$$

Доведення.

$$\begin{aligned} \text{Cov}(X, Y) &= E(X - EX)(Y - EY) = EXY - E(X \cdot EY) - E(EX \cdot Y) + EX \cdot EY \\ &= EXY - EX \cdot EY \end{aligned}$$

$$\boxed{2.} \quad \text{Cov}(X, X) = DX$$

Доведення. Згідно властивості 1

$$\text{Cov}(X, X) = EX^2 - (EX)^2 = DX$$

[3.] Якщо X і Y – незалежні, то $Cov(X, Y) = 0$

Доведення. Оскільки X і Y незалежні, то $EXY = EX \cdot EY$. Тоді згідно властивості 1

$$Cov(X, Y) = EXY - EX \cdot EY = EX \cdot EY - EX \cdot EY = 0.$$

Коефіцієнт кореляції

Визначення. Коефіцієнтом кореляції випадкових величин X і Y називається

$$r(X, Y) \stackrel{\text{def}}{=} \frac{Cov(X, Y)}{\sqrt{DX}\sqrt{DY}} = \frac{E(X - EX)(Y - EY)}{\sqrt{DX}\sqrt{DY}}$$

Властивості коефіцієнта кореляції

[1.] $r(X, X) = 1$

[2.] $-1 \leq r(X, Y) \leq 1$

[3.] $r(X, Y) = \pm 1$ тоді і тільки тоді коли X і Y лінійно залежні

$$Y = aX + b$$

[4.] Якщо X і Y – незалежні, то $r(X, Y) = 0$

Доведення випливає з властивості 3 коваріації.

Випадкові величини X і Y некорельовані, якщо $r(X, Y) = 0$. Згідно властивості 4 незалежність випадкових величин тягне за собою некорельованість. Обернене загалом не має місця.

Кореляційна матриця

Нехай $\vec{X} = (X_1, X_2, \dots, X_n)$ випадковий вектор. Матриця

$$\{r(X_i, X_j): 1 \leq i \leq n, 1 \leq j \leq n\}$$

Кореляційною матрицею випадкового вектору $\vec{X}$.

Наприклад, якщо $\vec{X} = (X_1, X_2, X_3)$ - 3-вимірний випадковий вектор.

Позначимо $r_{ij} = r(X_i, X_j)$. Тоді

$$\begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix}$$

-кореляційна матриця цього вектору.

Задача лінійного прогнозу

Розглянемо дві випадкові величини X_1, X_2 . Припустимо, що підлягає спостереженню випадкова величина X_1 , а величина X_2 не спостерігається. Якщо величини X_1, X_2 корельовані, то є надія, що знання X_1 допоможе отримати деяку інформацію про X_2 . Розглянемо лінійну оцінку (лінійний прогноз) випадкової величини X_2 по спостереженням X_1 у вигляді

$$\tilde{X} = aX_1 + b.$$

Треба знайти такі $\hat{a}$ і $\hat{b}$, щоб лінійний прогноз $\tilde{X}$ був найкращим (оптимальним) в середньоквадратичному сенсі, тобто середньоквадратична похибки була б мінімальна

$$\Delta = E(\tilde{X} - X_2)^2 \rightarrow \min$$

Розглянемо випадковий вектор $\vec{X} = (X_1, X_2)$, з коваріаційною матрицею

$$\begin{pmatrix} \sigma_1^2 & r_{12}\sigma_1\sigma_2 \\ r_{21}\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$$

Джерело: Гіхман І.І., Скороход А.В., Ядренко М.Й. Теорія ймовірностей та математична статистика. 1988.

Київ: «Вища школа». 439 с.

і $EX_1 = m_1, EX_2 = m_2$. Обчислимо середньоквадратичну похибку (застосувавши властивість дисперсії $D(Z) = EZ^2 - (EZ)^2$)

$$\begin{aligned}\Delta &= E(\tilde{X} - X_2)^2 = D(\tilde{X} - X_2) + (E(\tilde{X} - X_2))^2 = \\ &= D\tilde{X} - 2 \cdot Cov(\tilde{X}, X_2) + DX_2 + (E(\tilde{X} - X_2))^2 = \\ &= \underbrace{D(aX_1 + b)}_{a^2\sigma_1^2} - 2 \cdot \underbrace{Cov(aX_1 + b, X_2)}_{ar_{12}\sigma_1\sigma_2} + \sigma_2^2 + \underbrace{(E(aX_1 + b - X_2))^2}_{(am_1 + b - m_2)^2} = \\ &= a^2\sigma_1^2 - 2 \cdot a \cdot r_{12}\sigma_1\sigma_2 + \sigma_2^2 + (am_1 + b - m_2)^2\end{aligned}$$

За рахунок вибору b можна мінімізувати останній доданок, зробивши його рівним нулю ($am_1 + b - m_2 = 0$):

$$b = m_2 - am_1$$

Перші три доданки утворюють квадратичний тричлен від a . Мінімум він досягає у вершині параболи:

$$\hat{a} = -\frac{-2 \cdot r_{12}\sigma_1\sigma_2}{2\sigma_1^2} = r_{12} \frac{\sigma_2}{\sigma_1}$$

При значеннях параметрів

$$\begin{cases} \hat{a} = r_{12} \frac{\sigma_2}{\sigma_1} \\ \hat{b} = m_2 - m_1 \cdot r_{12} \frac{\sigma_2}{\sigma_1} \end{cases} \quad (*)$$

Мінімальна похибка дорівнює

$$\Delta = \left(r_{12} \frac{\sigma_2}{\sigma_1}\right)^2 \sigma_1^2 - 2 \cdot r_{12} \frac{\sigma_2}{\sigma_1} \cdot r_{12}\sigma_1\sigma_2 + \sigma_2^2 = \sigma_2^2(1 - r_{12}^2)$$

Якщо $|r_{12}|=1$, то похибки немає. В той же час, чим ближче r_{12} до нуля – тим більше похибка. Лінійна апроксимація величини X_2 , що не спостерігається через значення величини X_1 , що спостерігається (тобто оптимальний лінійний прогноз), має вигляд

$$\tilde{X} = r_{12} \frac{\sigma_2}{\sigma_1} \cdot X_1 + m_2 - m_1 \cdot r_{12} \frac{\sigma_2}{\sigma_1} = r_{12} \frac{\sigma_2}{\sigma_1} \cdot (X_1 - m_1) + m_2$$

Простір випадкових величин (з дисперсіями). Розглянемо множину центрованих випадкових величин (з нульовими математичними сподіваннями). Якщо X – випадкова величина з цієї множини, то для любого числа α випадкова величина $\alpha \cdot X$ теж належить цій множині. Якщо X і Y – випадкові величини з цієї множини, то їх сума $X + Y$ теж належить цій множині. Тобто такі величини утворюють лінійний простір і тому кожен випадкову величину ми можемо уявляти як вектор. В цьому лінійному просторі можна ввести скалярний добуток

$$\langle X, Y \rangle \stackrel{\text{def}}{=} \text{Cov}(X, Y)$$

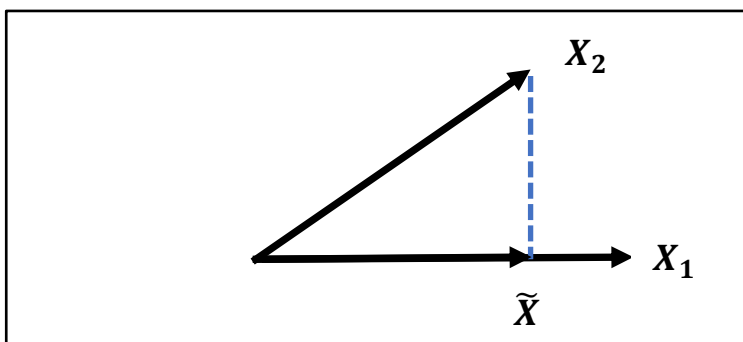
Тоді довжина вектору

$$\|X\| \stackrel{\text{def}}{=} \sqrt{\langle X, X \rangle}$$

Та косинус кута між векторами

$$\cos(X, Y) = \frac{\langle X, Y \rangle}{\|X\| \|Y\|} = \frac{\text{Cov}(X, Y)}{\sqrt{DX} \sqrt{DY}} = r_{XY}$$

- це коефіцієнт кореляції. Зокрема некорельованість випадкових величин означає їх ортогональність ($\cos(X, Y) = 0 \Leftrightarrow X \perp Y$). Тоді лінійний прогноз може бути зображений наступним чином



РОЗДІЛ 3. ОСНОВИ СТАТИСТИЧНОГО АНАЛІЗУ ДАНИХ

ТЕМА 11. УЗАГАЛЬНЕННЯ ТА ПРЕДСТАВЛЕННЯ ДАНИХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 12

Математичною статистикою називається наука, яка займається розробкою методів відбору, опису та аналізу дослідних даних з метою вивчення закономірностей випадкових (часто масових) явищ. У свою чергу, встановлення цих закономірностей ґрунтується на вивченні методами теорії ймовірностей статистичних даних – результатів дослідів або спостережень. Отже, математична статистика вивчає методи, які дають змогу за результатами випробувань робити певні ймовірнісні висновки.

Основні задачі математичної статистики:

- вказати способи збору та групування (якщо даних дуже багато) статистичної інформації;
- визначити закон розподілу випадкової величини або системи випадкових величин за статистичними даними;
- оцінити невідомі параметри розподілу;
- перевірити правдоподібність припущень про закон розподілу випадкової величини, про форму зв'язку між випадковими величинами або про значення оцінюваного параметра.

Можна сказати, що основне завдання математичної статистики – **розроблення методів аналізу статистичних даних залежно від мети дослідження**.

Суцільне чи не суцільне спостереження ?

Зрозуміло, що у випадку суцільних опитувань (спостережень) ми будемо мати повне враження про характеристики генеральної сукупності, наприклад, цільової аудиторії (ЦА). Але в більшості життєвих ситуацій ми знаходимося в умовах обмеженості ресурсів: недостатньо грошей, обмежений час, не вистачає персоналу для виконання роботи по збору статистичних даних. Часто взагалі неможливо провести суцільне спостереження. Тому, природньо, що при проведенні економічних, соціологічних та маркетингових досліджень перевагу віддають несуцільним опитуванням.

Але, чи можна за частиною судити про ціле ? Чи можна довіряти цим висновкам? Наскільки велика похибка? Дійсно, за частиною можна судити про ціле. Якщо «правильно» все робити, то висновкам можна довіряти. Більш того, можна оцінити похибку наших узагальнень. Відповіді на питання як це зробити дають процедури і методи, які об'єднанні назвою – **вибіркові статистичні спостереження**.

Перифразуючи висловлювання відомого фахівця з соціологічних досліджень суспільної думки Джоржа Геллапа сформулюємо наступний принцип «українського борщу», який лежить в основі випадкових вибірових спостережень:

Щоб дізнатися смак борщу - не треба їсти всю каструлю. Досить добре перемішати борщ і спробувати одну ложку.

Процедура «перемішування» борщу гарантує, що смак борщу в ложці буде представляти (*презентувати*) смак борщу в каструлі. Так само в вибірових дослідженнях деяка процедура «перемішування» при формуванні вибірової сукупності (***вибірки***) з ***генеральної сукупності*** (кажемо: досліджуваної множини, наприклад, цільової аудиторії) гарантує, що концентрація ознаки (наприклад, рейтинг політичного діяча) в вибірці приблизно буде такою, як в генеральній сукупності. Тобто вибірка «добре» презентує генеральну сукупність, а отже можна казати про її **репрезентативність**.

Поняття репрезентативність вибірки відносно. Одна і та ж вибірка (звісно за умов правильного «перемішування») може бути репрезентативною для однієї аудиторії і не бути репрезентативною для іншої.

CASE. Широко відомий випадок електоральних перегонів кандидатів у президенти США в 1936 році за участю Теодора Рузвельта та Альфреда Лендона. Американський журнал «Літературний огляд» напередодні виборів керуючись випадковим вибором розіслав більше 2 мільйонів поштових листівок з проханням назвати ім'я майбутнього президента. Згідно результатам опитування, президентом було прогнозовано А. Лендона. В той же час Джордж Геллап базуючи свої висновки на вибірці об'ємом тільки в 4000 анкет спрогнозував перемогу Т. Рузвельту. Як відомо президентом США став Т. Рузвельт.

Два прогнози... Один базується на виборці розміром в 4000, а другий в 2 мільйони. В першому випадку затрачено набагато менше ресурсів, а результат – кращий. В чому причина хибного прогнозу? Випадковий вибір, який був застосований для відбору респондентів – дуже добре, бо відповідає ідеології вибірових спостережень !!! Але вибір проводився за телефонними книгами. В той час в нього входили найбільш забезпечені верстви американського суспільства, зокрема - домовласники. Ця вибірка була репрезентативна для соціуму домовласників, але не була репрезентативна для всіх виборців США. Тому і такий результат. Мабуть, тому і з'явився жарт, що є «правда – маленька брехня – брехня – велика брехня – статистика».

Методи вибіркового відбору

Детерміновані методи відбору:

- квотна
- «першого зустрічного»
- «снігової кулі»
- поверхнева (експертна)
- «вибірка добровольців»
-

Ймовірнісні методи відбору:

- проста випадкова
- систематична
- стратифікована
- кластерна (гніздовий)
- стратифікована
- кластерна

Вибірка називається *ймовірнісною*, якщо кожен представник генеральної сукупності має відому ненульову ймовірність бути відібраним.

Вибірка називається *детермінованою*, якщо неможливо обчислити вірогідність відбору кожного представника генеральної сукупності або якась частина представників генеральної сукупності взагалі не має можливість потрапити до вибірки.

Модель ймовірнісної вибірки

Теоретична компонента. Вважаємо, що процедура відбору до вибірки забезпечує незалежність залучення до вибірки кожного спостереження. З теоретичної точки зору можемо розглядати вибірку обсягом n

$$x_1, x_2, \dots, x_n,$$

як конкретну реалізацію незалежних однаково розподілених випадкових величин

$$X_1, X_2, \dots, X_n,$$

що розподілені як досліджувана характеристика (випадкова величина) X з функцією розподілу $F(t)$. Тобто X описує розподіл досліджуваної характеристики в генеральній сукупності, а $F(t)$ - є її функцією розподілу

Випадкова величина X називається теоретичною випадковою величиною, а її функція розподілу $F(t)$ називається теоретичною функцією

розподілу. Вони є об'єктами дослідження математичної статистики і підлягають всебічній пильній увазі дослідника.

Емпірична компонента. Дані для аналізу (та оцінки X і $F(t)$) ми отримуємо з дослідів, тобто з статистичних спостережень. В реаліях маємо n спостережень

$$x_1, x_2, \dots, x_n, \quad (*)$$

за допомогою яких треба оцінити розподіл досліджуваної характеристики (тобто теоретичної випадкової величини) $F(t)$.

Визначимо допоміжну випадкову величину, за допомогою якої будемо оцінювати теоретичний розподіл. Оскільки ми не знаємо теоретичний розподіл, то кожне значення, отримане у вибірці, розглядаємо як таке, що може з'явитися з однаковою ймовірністю. Випадкову величину $\hat{X}$, що задається відповідним законом розподілу будемо називати **емпіричною випадковою величиною**:

$$x_1, x_2, \dots, x_n, \quad (*)$$

за допомогою яких треба оцінити розподіл досліджуваної характеристики (тобто теоретичної випадкової величини) $F(t)$.

$\hat{X}$	x_1	x_2	$\dots$	x_n
P	$\frac{1}{n}$	$\frac{1}{n}$	$\dots$	$\frac{1}{n}$

Характеристики емпіричної випадкової величини називають емпіричними, зокрема, її функцію розподілу $\hat{F}(t)$ – емпіричною функцією розподілу. Формально її можна записати у вигляді (вважаємо, що значення випадкової величини $\hat{X}$ впорядковані $x_1 < x_2 < \dots < x_n$):

$$\hat{F}(t) = \begin{cases} 0, & \text{якщо } t \leq x_1 \\ \frac{k}{n}, & \text{якщо } x_k \leq t < x_{k+1} \\ 1, & \text{якщо } t > x_n \end{cases}$$

Задача математичної статистики: за допомогою емпіричних даних (емпіричної функції розподілу $\hat{F}(t)$) оцінити реальний розподіл досліджуваної характеристики (теоретичну функцію розподілу $F(t)$).

Полігон і гістограма

Отримані за допомогою спеціальних досліджень (спостережень) дані надходять до дослідника у вигляді неструктурованого масиву даних. Тому постає питання щодо оброблення і впорядкування даних.

Значення x_i вибірки називають варіантами. Послідовність варіант, розміщених у порядку зростання (спадання), називають варіаційним рядом. Якщо при цьому x_i повторюється n_i разів

$$(n_1 + n_2 + \dots + n_k = n)$$

то n_i називають **абсолютною частотою варіанти** x_i , а $\frac{n_i}{n}$ – відносною частотою варіанти x_i .

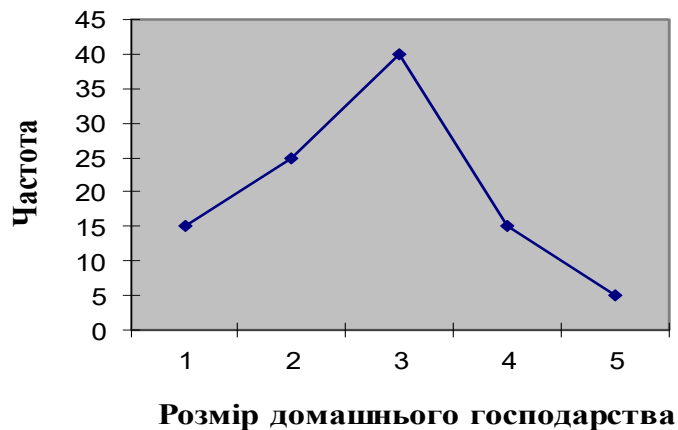
Визначення. Статистичним розподілом вибірки називають перелік варіант та відповідних частот або відносних частот. Статистичний закон розподілу зручно задавати таблицею, що встановлює зв'язок між значеннями випадкової величини та їх частотами:

$\hat{X}$	x_1	x_2	...	x_k
P	$\frac{n_1}{n}$	$\frac{n_2}{n}$	...	$\frac{n_k}{n}$

У випадку кратності (повторюваності) значень x_i емпірична функція розподілу приймає вигляд

$$\hat{F}(t) = \begin{cases} 0, & \text{якщо } t \leq x_1 \\ \sum_{i=1}^k \frac{n_i}{n}, & \text{якщо } x_k \leq t < x_{k+1} \\ 1, & \text{якщо } t > x_n \end{cases}$$

Визначення. *Полігоном частот* вибірки називають ламану з вершинами в точках (x_i, n_i) , де $i = 1, 2, \dots, k$. *Полігоном відносних частот* вибірки називають ламану з вершинами в точках $(x_i, \frac{n_i}{n})$, де $i = 1, 2, \dots, k$.



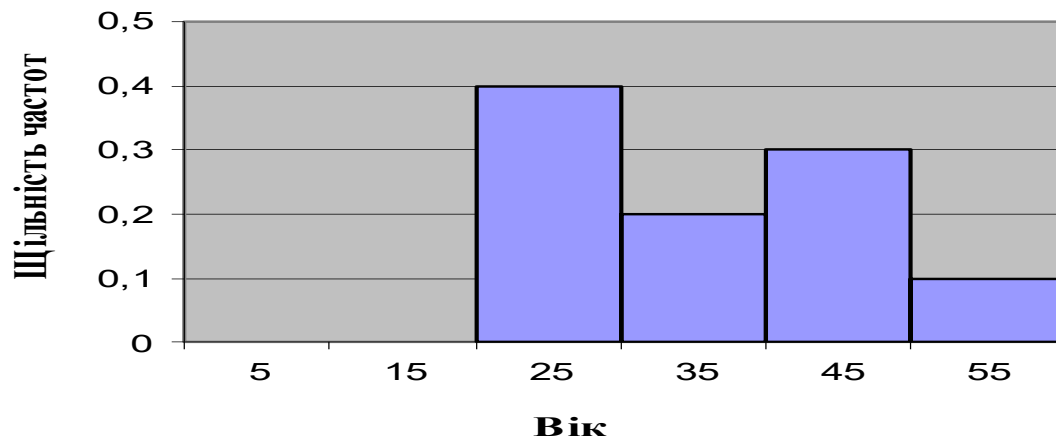
Якщо X – неперервна випадкова величина, то її статистичний розподіл також подають у вигляді таблиць. Для оцінювання її щільності за вибіркою

$$x_1, x_2, \dots, x_n$$

розбивають інтервал на частинні інтервали довжини h , у таблицю записують середину i -го інтервалу, n_i – кількість елементів вибірки i -го інтервалу.

Визначення Прямокутники з основами h і висотами $\frac{n_i}{nh}$ у прямокутній системі координат утворюють фігуру, яку називають *гістограмою вибірки*.

Часткові інтервали	Середини інтервалів	Частоти	Щільності частот
20-30	25	4	$4/(30-20) = 0,4$
30-40	35	2	$2/(40-30) = 0,2$
40-50	45	3	$3/(50-40) = 0,3$
50-60	55	1	$1/(60-50) = 0,1$



Точкові оцінки параметрів розподілу генеральної сукупності

Іноді з певних міркувань апіорі вдається встановити тип закону розподілу (наприклад, дихотомічний, нормальний, біноміальний, тощо) теоретичної випадкової величини, тобто розподілу характеристики генеральної сукупності, що досліджується. Тоді постає задача оцінювання параметрів цього розподілу.

Нехай X - теоретична випадкова величина з теоретичною функцією розподілу $F_{\theta}(t)$, де θ - невідомий параметр, що підлягає оцінюванню. В якості параметру θ можемо розглядати декілька невідомих параметрів

$$\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_k).$$

Визначення. Точковою оцінкою $\hat{\theta}$ невідомого параметру θ називається функція

$$\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$$

вибіркових значень (спостережень) $x_1, x_2, \dots, x_n$.

Визначення. Оцінка $\hat{\theta}$ параметра θ називається **незміщеною**, якщо

$$E\hat{\theta} = \theta \quad (*)$$

Оцінка $\hat{\theta}$ змінюється від вибірки до вибірки, тобто носить стохастичний (випадковий) характер, а параметр θ є постійною величиною оскільки описує закон розподілу досліджуваної характеристики генеральної сукупності.

Похибка оцінювання (наближення)

$$\Delta = \hat{\theta} - \theta$$

відповідно до визначення незміщеності (*) має ненульове математичне сподівання

$$E\Delta = 0,$$

що в середньому відхилення оцінки від невідомого значення параметру дорівнює нулю, тобто додатні відхилення компенсуються від'ємними відхиленнями, а самі відхилення пояснюються випадковістю відбору елементів вибірки. Незсувні оцінки описують *несистематичні похибки*.

Якщо

$$E\hat{\theta} = \theta + a \quad (**)$$

то a називають зсувом оцінки $\hat{\theta}$.

Співвідношення (**) означає, в незалежності від вибірки оцінка $\hat{\theta}$ буде в середньому відрізнятися від істинного значення параметра θ на величину a .

Такі оцінки описують *систематичні помилки*. Отже:

Н.В. *Незміщенність оцінки означає відсутність систематичної помилки.*

Точкова оцінка генерального середнього (математичного сподівання теоретичної випадкової величини)

Нехай X - теоретична випадкова величина з невідомим математичним сподіванням $\theta = EX$. Розглянемо емпіричну випадкову величину $\hat{X}$, що за визначенням задається законом розподілу

$\hat{X}$	x_1	x_2	...	x_n
P	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$

де $x_1, x_2, \dots, x_n$ - конкретна вибірка (деяка реалізація послідовності незалежних однаково з X розподілених випадкових величин). Емпіричним математичним сподіванням емпіричної випадкової величини $\hat{X}$ буде вираз

$$E\hat{X} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (***)$$

В якості оцінки $\hat{\theta}$ теоретичного математичного сподівання $\theta = EX$ візьмем середнє вибіркве (тобто емпіричне математичне сподівання $E\hat{X}$):

$$\hat{\theta} = \frac{X_1 + X_2 + \dots + X_n}{n} \quad (****)$$

де $X_1, X_2, \dots, X_n$ - вибірка розміру n , тобто послідовність незалежних однаково розподілених з X . Звідси

$$\theta = EX = EX_1 = EX_2 = \dots = EX_n$$

Підрахуємо математичне сподівання оцінки $\hat{\theta}$:

$$E\hat{\theta} = \frac{EX_1 + EX_2 + \dots + EX_n}{n} = \frac{\overbrace{\theta + \theta + \dots + \theta}^n}{n} = \theta$$

Згідно з визначенням оцінка $\hat{\theta}$ (****) є незміщеною.

N.B. Вибіркове середнє

$\hat{\theta} = \frac{X_1 + X_2 + \dots + X_n}{n}$
--

є незміщеною оцінкою генерального середнього.

**Точкова оцінка генеральної дисперсії
(дисперсії теоретичної випадкової величини)**

Нехай X - теоретична випадкова величина з невідомим математичним сподіванням $\theta = DX$. Для емпіричної випадкової величини $\hat{X}$, обчислимо дисперсію

$$D\hat{X} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

де $\bar{x} = E\hat{X}$ - вибіркове середнє, що визначається співвідношенням (***)

В якості оцінки $\hat{\theta}$ теоретичної дисперсії $\theta = DX$ візьмем вибірккову дисперсію (тобто емпіричну дисперсію $D\hat{X}$):

$$\hat{\theta} = D\hat{X} = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}$$

де $X_1, X_2, \dots, X_n$ - вибірка розміру n , тобто послідовність незалежних однаково розподілених з X випадкових величин,

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Знайдемо математичне сподівання оцінки $\hat{\theta}$:

$$\begin{aligned} E\hat{\theta} &= E\left(\frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}\right) = \\ &= \frac{n}{n-1} DX \end{aligned}$$

Виявилося, що

$$E\hat{\theta} = \frac{n}{n-1} DX = \frac{n}{n-1} \theta \neq \theta$$

Отже оцінка генеральної дисперсії $\theta = DX$ через значення вибіркової дисперсії не є незміщеною. Однак, якщо трохи «підправити» вибірккову дисперсію таким чином

$$\hat{\theta} = \frac{n}{n-1} D\hat{X} \tag{v}$$

то отримаємо

$$E \hat{\theta} = \frac{n}{n-1} DX = \theta$$

Це означає, що оцінка (v) буде незміщеною. Вона відрізняється від вибіркової дисперсії на множник $\frac{n}{n-1}$.

Н.В. Виправлена вибіркова дисперсія

$$\hat{\theta} = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1}$$

є незміщеною оцінкою генеральної дисперсії.

Визначення. Послідовність оцінок

$$\hat{\theta}_n \equiv \hat{\theta}_n(x_1, x_2, \dots, x_n)$$

Називається **слушною**, якщо вона збігається за ймовірністю до невідомого параметру θ :

$$\hat{\theta}_n \xrightarrow{P} \theta, \quad n \rightarrow \infty$$

Визначення. Послідовність оцінок

$$\hat{\theta}_n \equiv \hat{\theta}_n(x_1, x_2, \dots, x_n)$$

Називається **сильно слушною**, якщо вона збігається з ймовірністю одиниця до невідомого параметру θ :

$$P \left(\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta \right) = 1.$$

Точкова оцінка генерального середнього (математичного сподівання теоретичної випадкової величини)

Нехай X - теоретична випадкова величина з невідомим математичним сподіванням $\theta = EX$. В якості оцінки $\hat{\theta}$ теоретичного математичного сподівання $\theta = EX$ візьмем середнє вибіркве (тобто емпіричне математичне сподівання $E\hat{X}$):

$$\hat{\theta} = \frac{X_1 + X_2 + \dots + X_n}{n} \quad (*)$$

де $X_1, X_2, \dots, X_n$ - вибірка розміру n , тобто послідовність незалежних однаково розподілених з X .

Згідно закону великих чисел

$$\hat{\theta}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{P} \theta = EX, \quad n \rightarrow \infty$$

Згідно визначенню послідовність оцінок $\hat{\theta}_n$ є слушною. Більш того ця послідовність оцінок буде сильно слушною в силу посиленого закону великих чисел

$$P \left(\hat{\theta}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{n \rightarrow \infty} \theta = EX \right) = 1$$

N.B. *Послідовність вибірових середніх*

$$\hat{\theta}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

є (сильно) слушною оцінкою генерального середнього

Визначення. Нехай $\hat{\theta}_1$ и $\hat{\theta}_2$ оцінки невідомого параметру θ .

Кажуть, що оцінка $\hat{\theta}_1$ ефективніше ніж оцінка $\hat{\theta}_2$, якщо

$$D\hat{\theta}_1 \leq D\hat{\theta}_2$$

ТЕМА 12. ТОЧКОВЕ ОЦІНЮВАННЯ ПАРАМЕТРІВ СТОХАСТИЧНИХ МОДЕЛЕЙ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 13

Методи точкового оцінювання параметрів розподілу генеральної сукупності

Метод моментів

Суть методу моментів полягає в прирівнюванні теоретичних моментів до емпіричних, а рішення відповідних рівнянь називаються оцінками параметрів за методом моментів.

Нехай X - теоретична випадкова величина, а $\hat{X}$ - емпірична випадкова величина. І, нехай невідомими є k моментів $\theta_1, \theta_2, \dots, \theta_k$. Розглянемо перші k теоретичні моменти:

$$EX = a_1(\theta_1, \theta_2, \dots, \theta_k)$$

$$EX^2 = a_2(\theta_1, \theta_2, \dots, \theta_k)$$

...

$$EX^k = a_k(\theta_1, \theta_2, \dots, \theta_k)$$

Та перші k емпіричні моменти

$$\begin{cases} E\hat{X} = \widehat{a_1} \\ E\hat{X}^2 = \widehat{a_2} \\ \dots \\ E\hat{X}^k = \widehat{a_k} \end{cases}$$

Рішення системи рівнянь

$$\begin{cases} a_1(\theta_1, \theta_2, \dots, \theta_k) = \widehat{a_1} \\ a_2(\theta_1, \theta_2, \dots, \theta_k) = \widehat{a_2} \\ \dots \\ a_k(\theta_1, \theta_2, \dots, \theta_k) = \widehat{a_k} \end{cases}$$

Будуть оцінками невідомих параметрів $\theta_1, \theta_2, \dots, \theta_k$ за методом моментів.

Приклад 1. Нехай теоретичний розподіл X є дихотомічним

X	0	1
P	$q = 1 - \theta$	θ

де θ - невідомий параметр, що підлягає оцінюванню. Застосуємо метод моментів оцінювання параметру θ .

Перший теоретичний момент

$$EX = 0 \cdot (1 - \theta) + 1 \cdot \theta = \theta$$

Якщо $X_1, X_2, \dots, X_n$ - вибірка, то емпіричний перший момент має вигляд

$$E\hat{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Тоді згідно процедури методу моментів прирівняємо теоретичний і емпіричний моменти і отримаємо рішення

$$\hat{\theta} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

що і є оцінкою параметра за методом моментів.

Приклад 2. Нехай теоретична випадкова величина X має нормальний розподіл с параметрами θ_1 та θ_2 (Позначаємо як $N(\theta_1, \theta_2^2)$). Як відомо $\theta_1 = EX$, $\theta_2 = DX$, а щільність має вигляд

$$P(u) = \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(u-\theta_1)^2}{2\theta_2}}$$

Перший і другий початкові теоретичні моменти:

$$EX, \quad EX^2 = DX + (EX)^2$$

(останнє випливає з властивості дисперсії $DX = EX^2 - (EX)^2$)

Перший і другий початкові емпіричні моменти

$$E\widehat{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$E\widehat{X}^2 = \frac{X_1^2 + X_2^2 + \dots + X_n^2}{n}$$

Прирівнюючи теоретичні та емпіричні моменти отримуємо систему рівнянь

$$\begin{cases} \theta_1 = \frac{X_1 + X_2 + \dots + X_n}{n} \\ \theta_2 + \theta_1^2 = \frac{X_1^2 + X_2^2 + \dots + X_n^2}{n} \end{cases}$$

Розв'язавши цю систему отримуємо рішення, що є оцінками параметрів за методом моментів

$$\begin{cases} \widehat{\theta}_1 = \frac{X_1 + X_2 + \dots + X_n}{n} \equiv \bar{X} \\ \widehat{\theta}_2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n} \end{cases}$$

Метод максимальної правдоподібності

Суть методу максимальної правдоподібності полягає в тому, що за оцінку $\hat{\theta}$ невідомого параметру беруть таке значення θ , для якого ймовірність отримання вже наявної вибірки максимальна.

Визначення. Функцією правдоподібності дискретної випадкової величини X називають функцію аргументу θ :

$$L(x_1, x_2, \dots, x_n, \theta) = p(x_1, \theta) \cdot p(x_2, \theta) \dots p(x_n, \theta)$$

де $x_1, x_2, \dots, x_n$ — вибіркові значення, а

$$p(x_i, \theta) = P(X = x_i)$$

Приклад. Нехай теоретичний розподіл X є дихотомічним

X	0	1
P	$q = 1 - \theta$	θ

де θ - невідомий параметр, що підлягає оцінюванню. Нехай $x_1, x_2, \dots, x_n$ – деяка наявна вибірка. Оскільки теоретична випадкова величина приймає тільки два значення 0 і 1, то вибірка складається з нулів та одиниць. Припустимо, у вибірці k одиниць (відповідно нулів буде $n-k$). Щоб застосувати метод максимальної правдоподібності побудуємо функцію правдоподібності

$$L(\theta) = \theta^k (1 - \theta)^{n-k}$$

Для знаходження максимуму функції $L(\theta)$ візьмемо похідну і прирівняємо її до нуля:

$$\begin{aligned} L'(\theta) &= k\theta^{k-1}(1 - \theta)^{n-k} - (n - k)\theta^k(1 - \theta)^{n-k-1} = \\ &= \theta^{k-1}(1 - \theta)^{n-k-1}(k(1 - \theta) - (n - k)\theta) = \\ &= \theta^{k-1}(1 - \theta)^{n-k-1}(k - n\theta) = 0 \quad \Rightarrow \\ (k - n\theta) &= 0 \quad \Leftrightarrow \quad \theta = \frac{k}{n} \end{aligned}$$

В точці

$$\theta = \frac{k}{n}$$

функція $L(\theta)$ має максимум, тобто це значення є оцінкою параметру за методом максимальної правдоподібності, яке можна представити у вигляді

$$\hat{\theta} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Визначення. Функцією правдоподібності неперервної випадкової величини X називають функцію аргументу θ :

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \dots f(x_n, \theta)$$

де $x_1, x_2, \dots, x_n$ – вибіркові значення, а $f(x_i, \theta)$ – щільність розподілу.

Приклад. Нехай теоретична випадкова величина X має нормальний розподіл с параметрами θ_1 та θ_2 . Відомо $\theta_1 = EX$, $\theta_2 = DX$, а щільність має вигляд

$$f(u) = \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(u-\theta_1)^2}{2\theta_2}}$$

Якщо $x_1, x_2, \dots, x_n$ – вибіркові значення, то функція правдоподібності набирає вигляду

$$\begin{aligned} L(\theta_1, \theta_2) &= \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_1-\theta_1)^2}{2\theta_2}} \cdot \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_2-\theta_1)^2}{2\theta_2}} \cdot \dots \cdot \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_n-\theta_1)^2}{2\theta_2}} = \\ &= (2\pi\theta_2)^{-n/2} \cdot e^{-\frac{\sum_{i=1}^n (x_i-\theta_1)^2}{2\theta_2}} \end{aligned}$$

Для того, щоб знайти максимум функції треба знайти часткові похідні і прирівняти їх до нуля. Зауважимо, що $L(\theta_1, \theta_2)$ і $\ln L(\theta_1, \theta_2)$ досягає максимуму в одній точці. Тому з міркувань простоти будемо шукати максимум функції $l(\theta_1, \theta_2) = \ln L(\theta_1, \theta_2)$:

$$\left\{ \begin{array}{l} \frac{\partial l(\theta_1, \theta_2)}{\partial \theta_1} = 0 \\ \frac{\partial l(\theta_1, \theta_2)}{\partial \theta_2} = 0 \end{array} \right. \Leftrightarrow$$

$$\begin{cases} \sum_{i=1}^n x_i - n \theta_1 = 0 \\ -n\theta_2 + \sum_{i=1}^n (x_i - \theta_1)^2 = 0 \end{cases}$$

Звідси

$$\begin{cases} \hat{\theta}_1 = \frac{X_1 + X_2 + \dots + X_n}{n} \equiv \bar{X} \\ \hat{\theta}_2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n} \end{cases}$$

Метод максимуму правдоподібності оцінювання параметрів розподілу має низку важливих переваг: вони слушні, асимптотичне нормальні (за великих n їх розподіл близький до нормального) і мають найменшу дисперсію порівняно з іншими асимптотичне нормальними оцінками. Цей метод найбільш повно використовує дані вибірки для оцінювання параметрів, тому він особливо корисний для малих обсягів вибірки. Однак його недолік полягає у тому, що він вимагає складних обчислень.

Поняття та розподіли, що використовуються при реалізації техніки статистичних обчислень

Визначення. *Варіаційним рядом* числової послідовності

$$x_1, x_2, \dots, x_n$$

називається перестановка членів цієї послідовності (які позначаються $x_{(1)}, x_{(2)}, \dots, x_{(n)}$) у зростаючому порядку

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Визначення. *Квантилем* порядку q ($0 < q < 1$), або *q-квантилем* випадкової величини X називається таке число t_q , що

$$P(X \leq t_q) \geq q,$$

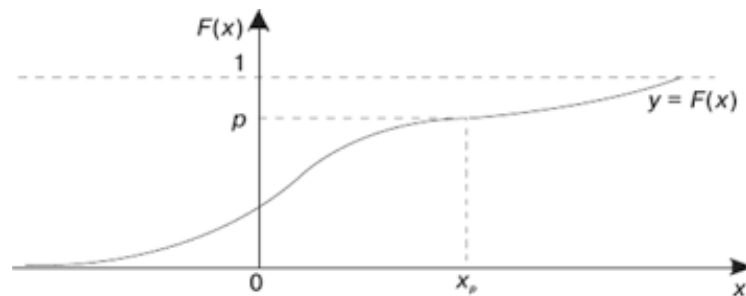
$$P(X \geq t_q) \geq 1 - q$$

Якщо ймовірність задана у відсотках, квантиль називається **процентилем** або **перцентилем**. Квантиль порядку $\frac{1}{2}$ називається **медіаною**.

Якщо розподіл безперервний, то **q**-квантиль однозначно задається рівнянням

$$F(t_q) = q$$

де $F(t)$ – функція розподілу випадкової величини X .



Визначення. Випадкова величина

$$Y = f(X_1, X_2, \dots, X_n)$$

що є функцією від вибірових спостережень

$$X_1, X_2, \dots, X_n$$

називається статистикою (у вузькому значенні).

Нормальний розподіл

Визначення. Випадкова величина X нормально розподілена $N(a, \sigma^2)$, якщо щільність має вид

$$P(u) = \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(u-\theta_1)^2}{2\theta_2}}$$

(де $a = EX$ – математичне сподівання, а $\sigma^2 = DX$ – дисперсія)

Оцінка генерального середнього

$$\hat{\theta} = \frac{\bar{X} - a}{\sigma/\sqrt{n}}$$

де

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{S_n}{n}$$

- вибіркове середнє.

Для того, щоб мати уявлення про якість оцінки, необхідно знати розподіл цієї оцінки. Розглянемо це питання у випадку нормально розподіленої генеральної сукупності. Будемо припускати, що відома дисперсія σ^2 цього розподілу.

Теорема 1. Нехай $X_1, X_2, \dots, X_n$ – вибірка з нормального розподілу (послідовність незалежних нормально розподілених випадкових величин), тоді статистика (випадкова величина)

$$S_n = X_1 + X_2 + \dots + X_n$$

має нормальний розподіл.

Теорема 2. Нехай $X_1, X_2, \dots, X_n$ – вибірка з нормального $N(a, \sigma^2)$ - розподілу (послідовність незалежних $N(a, \sigma^2)$ - розподілених випадкових величин), тоді статистика (випадкова величина)

$$N = \frac{\bar{X} - a}{\sigma/\sqrt{n}}$$

має стандартний нормальний розподіл (тобто $N(0,1)$ -розподіл), де

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{S_n}{n}$$

-вибіркове середнє.

Теорема 3. Нехай $X_1, X_2, \dots, X_n$ – вибірка (не обов'язково з нормального розподілу), тоді статистика (випадкова величина)

$$Y = \frac{\bar{X} - a}{\sigma/\sqrt{n}}$$

наближається до стандартного нормального розподілу (тобто $N(0,1)$ -розподілу), де

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{S_n}{n}$$

-вибіркове середнє.

Тобто для «великих» виборок ($n \geq 50$, іноді $n \geq 30$) статистика (випадкова величина)

$$Y = \frac{\bar{X} - a}{\sigma/\sqrt{n}}$$

приблизно має стандартний нормальний розподіл

На «кандидата» оцінки генерального середнього у випадку невідомої дисперсії σ^2 може слугувати наступна оцінка

$$\hat{\theta} = \frac{\bar{X} - a}{\hat{\sigma}/\sqrt{n}} \quad (*)$$

де $\bar{X}$ - вибіркове середнє, $\hat{\sigma}^2$ – вибіркова дисперсія.

На жаль, розподіл оцінки $(*)$ не є нормальним. Як виявилось розподіл цієї оцінки є розподілом Стюдента.

Розподіл Стюдента (Student's distribution)

Розподіл Стюдента (Додаток А) було запроваджено 1908 року В.С.Госсетом, ірландським службовцем пивоварного заводу, який брав участь у розробці нових технологій виробництва пива та жодним студентом не був. Додавати відомості результати досліджень означало відкрити корпоративну таємницю, тому В.С.Госсет надрукував свої матеріали під псевдонімом Стюдент. Фішер увів для нього позначення t-розподіл.

Визначення. Нехай $X_0, X_1, X_2, \dots, X_n$ – послідовність незалежних $N(a, \sigma^2)$ нормально розподілених випадкових величин

Тоді випадкова величина

$$t(k) = \frac{X_0}{\sqrt{\frac{1}{k} (X_1^2 + X_2^2 + \dots + X_k^2)}} = \frac{X_0}{\sqrt{\frac{1}{k} \chi^2(k)}}$$

має розподіл Стюденту з k ступенями свободи.

Для оцінки генеральної дисперсії розглянемо наступну оцінку

$$\frac{s^2}{\sigma^2} = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{\sigma^2 \cdot (n-1)} \quad (**)$$

Виявилось, що вона має хі-квадрат розподіл.

Теорема 4. Нехай $X_1, X_2, \dots, X_n$ – вибірка з нормального $N(a, \sigma^2)$ -розподілу (послідовність незалежних $N(a, \sigma^2)$ - розподілених випадкових величин), тоді статистика (випадкова величина)

$$t(n-1) = \frac{\bar{X} - a}{s/\sqrt{n}}$$

має **розподіл Стюдента з $n-1$ ступенями свободи**.

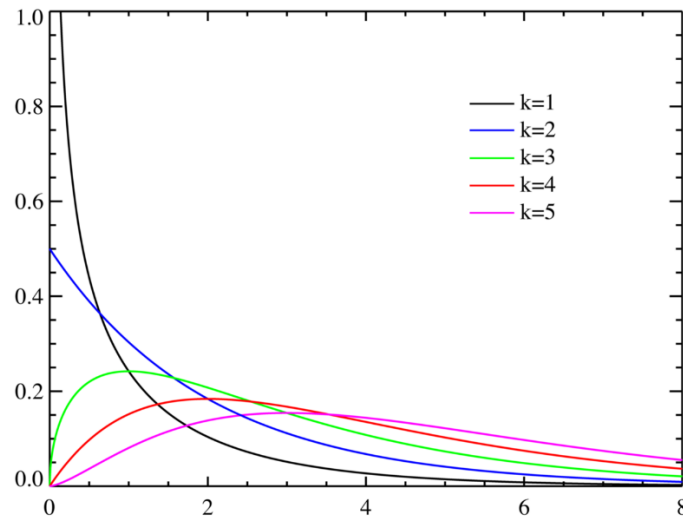
($\bar{X}$ - вибіркове середнє, s^2 – «підправлена» вибіркова дисперсія).

Хі-квадрат розподіл (χ^2 – розподіл)

Визначення. Нехай $X_1, X_2, \dots, X_n$ – послідовність незалежних нормально розподілених випадкових величин $N(0,1)$. Тоді випадкова величина

$$\chi^2(k) = X_1^2 + X_2^2 + \dots + X_k^2$$

має Хі-квадрат розподіл з k ступенями свободи (позначають $df=k$, degrees of freedom).



Джерело: Chan S.H. *Introduction to Probability for Data Science*. 2021. Michigan Publishing Services. 704 p.

Теорема 5. Нехай $X_1, X_2, \dots, X_n$ – вибірка з нормального $N(a, \sigma^2)$ -розподілу (послідовність незалежних $N(a, \sigma^2)$ - розподілених випадкових величин), тоді статистика (випадкова величина)

$$\frac{s^2}{\sigma^2} = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{\sigma^2 \cdot (n - 1)}$$

має **хі-квадрат розподіл з $n-1$ ступенями свободи**.

($\bar{X}$ - вибіркове середнє, s^2 – «підправлена» вибіркова дисперсія).

Нехай $X_1, X_2, \dots, X_k, X_{k+1}, X_{k+2}, \dots, X_{k+m}$ дві виборки з нормального $N(a, \sigma^2)$ -розподілу. Розглянемо дві статистики

$$V^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_k - \bar{X})^2}{(k - 1)}$$

$$U^2 = \frac{(X_{k+1} - \bar{X})^2 + (X_{k+2} - \bar{X})^2 + \dots + (X_{k+m} - \bar{X})^2}{(m - 1)}$$

Статистика

$$W^2 = \frac{V^2}{U^2}$$

застосовується при оцінці дисперсії генеральної сукупності і має розподіл Фішера (Додаток Б).

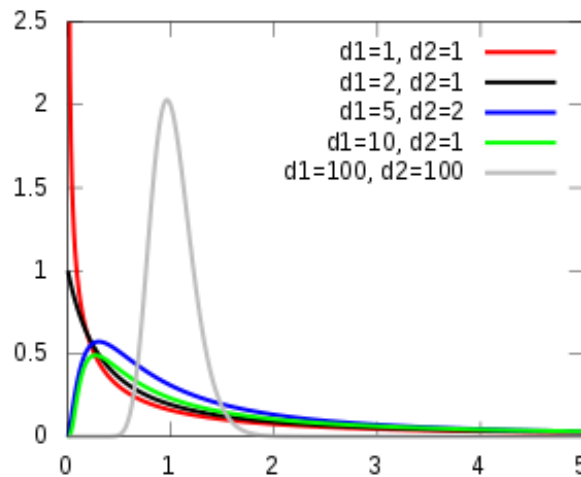
Розподіл Фішера (Fisher distribution, F-distribution)

Визначення. Нехай $X_1, X_2, \dots, X_k, X_{k+1}, X_{k+2}, \dots, X_{k+m}$ послідовність незалежних нормально розподілених випадкових величин $N(0, \sigma^2)$. Тоді випадкова величина

$$F(k, m) = \frac{\frac{1}{k} (X_1^2 + X_2^2 + \dots + X_k^2)}{\frac{1}{m} (X_{k+1}^2 + X_{k+2}^2 + \dots + X_{k+m}^2)} = \frac{\frac{1}{k} \chi^2(k)}{\frac{1}{m} \chi^2(m)}$$

має розподіл Фішера з k, m ступенями свободи.

Щільність розподілу Фішера з k, m ступенями свободи



Джерело: Chan S.H. *Introduction to Probability for Data Science*. 2021. Michigan Publishing Services. 704 p.

Властивість розподілу Фішера

$$F(k, m) = \frac{1}{F(m, k)}$$

ТЕМА 13. ІНТЕРВАЛЬНЕ ОЦІНЮВАННЯ ПАРАМЕТРІВ ТА ДОВІРЧІ ІНТЕРВАЛИ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 14

Довірчий інтервал

При точковій оцінці параметра θ оцінка $\hat{\theta}$ цього параметра є наближеним значенням невідомого параметра навіть у разі коли вона є незміщена, слушною та ефективною. Виникає питання: наскільки $\hat{\theta}$ може відхилятися від істинного значення θ . Чи можна вказати таку величину Δ , яка з практичною достовірністю гарантує виконання нерівності

$$|\theta - \hat{\theta}| \leq \Delta \quad ?$$

Тобто, чи можна вказати інтервал

$$(\hat{\theta}_1, \hat{\theta}_2)$$

який би з заздалегідь заданою ймовірністю (близькою до одиниці) накривав би невідомий параметр θ ?

Ймовірність (близькою до одиниці), що заздалегідь обирається дослідником для виконання нерівності

$$\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2$$

зазвичай називається довірчою ймовірністю, а сам інтервал

$$(\hat{\theta}_1, \hat{\theta}_2)$$

Називається довірчим інтервалом.

Визначення. Інтервал $(\hat{\theta}_1, \hat{\theta}_2)$ називається довірчим (confidence interval) з рівнем значущості α , якщо невідомий параметр θ накривається ним з ймовірністю $1 - \alpha$:

$$P(\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2) = 1 - \alpha$$

$1 - \alpha$ називається довірчою ймовірністю.

Рівень значущості – це ймовірність з якою істинне значення параметру θ не попадає в довірчий інтервал $(\hat{\theta}_1, \hat{\theta}_2)$.

Величини $\hat{\theta}_1, \hat{\theta}_2, \Delta$ будуються з урахуванням вибірових значень

$$x_1, x_2, \dots, x_n,$$

тобто є функціями від значень спостережень.

Ширина довірчого інтервалу істотно залежить від обсягу вибірки n (зменшується зі зростанням n) і від величини довірчої ймовірності $1 - \alpha$ (збільшується з наближенням довірчої ймовірності до одиниці).

Інтервальна оцінка для середнього нормального розподілу.

Випадок відомої дисперсії

Нехай X теоретична випадкова величина, що має нормальний розподіл

$$N(\theta, \sigma^2)$$

де $\theta = EX$ – невідоме математичне сподівання, а $\sigma^2 = DX$ – відома дисперсія. Нехай $X_1, X_2, \dots, X_n$ – вибірка, тобто послідовність незалежних нормально розподілених випадкових величин $N(\theta, \sigma^2)$. Позначимо

$$S_n = X_1 + X_2 + \dots + X_n, \quad n \geq 1$$

В якості оцінки математичного сподівання візьмемо вибірове середнє

$$\hat{\theta} = \frac{S_n}{n} \equiv \bar{X}$$

Розглянемо похибку оцінювання Δ :

$$|\hat{\theta} - \theta| \leq \Delta$$

Це можна переписати у вигляді

$$\hat{\theta} - \Delta < \theta < \hat{\theta} + \Delta$$

Тому будемо шукати таке Δ , щоб інтервал $\hat{\theta} - \Delta < \theta < \hat{\theta} + \Delta$ містив невідомий параметр θ з ймовірністю $1 - \alpha$.

$$P\{|\hat{\theta} - \theta| \leq \Delta\} = P\{\hat{\theta} - \Delta < \theta < \hat{\theta} + \Delta\} = 1 - \alpha \quad (*)$$

Згідно з визначенням такий інтервал буде довірчим з рівнем значущості α .

Тобто ліва межа довірчого інтервалу

$$\hat{\theta}_1 = \hat{\theta} - \Delta$$

А права межа довірчого інтервалу

$$\hat{\theta}_2 = \hat{\theta} + \Delta$$

Співвідношення (*) можна переписати у вигляді

$$\begin{aligned} P\{|\hat{\theta} - \theta| \leq \Delta\} &= P\left\{\left|\frac{S_n}{n} - \theta\right| \leq \Delta\right\} = \\ &= P\left\{-\Delta \leq \frac{S_n}{n} - \theta \leq \Delta\right\} = 1 - \alpha \quad \Leftrightarrow \\ P\left\{\frac{-\Delta}{\sigma/\sqrt{n}} \leq \frac{\frac{S_n}{n} - \theta}{\sigma/\sqrt{n}} \leq \frac{\Delta}{\sigma/\sqrt{n}}\right\} &= 1 - \alpha \end{aligned}$$

За теоремою 2 лекції 13 випадкова величина

$$N = \frac{\frac{S_n}{n} - \theta}{\sigma/\sqrt{n}} = \frac{\bar{X} - \theta}{\sigma/\sqrt{n}}$$

має стандартний нормальний розподіл (тобто $N(0, 1)$ -розподіл);

$$P\{-u < N < u\} = 1 - \alpha \quad (**)$$

де

$$u = \frac{\Delta}{\sigma/\sqrt{n}}$$

Якщо $\Phi(u)$ -функція $N(0, 1)$ -розподілу, (**) можна переписати у вигляді

$$P\{-u < N < u\} = \Phi(u) - \Phi(-u) = 1 - \alpha \quad (***)$$

Оскільки

$$\Phi(u) - \underbrace{\Phi(-u)}_{1-\Phi(u)} = 2\Phi(u) - 1$$

то згідно (***)

$$2\Phi(u) - 1 = 1 - \alpha \quad \Rightarrow \quad \Phi(u) = 1 - \alpha/2$$

Згідно визначення квантилю, останнє рівняння задовольняє квантиль стандартного нормального розподілу порядку $1 - \alpha/2$

$$\Phi(u_{1-\alpha/2}) = 1 - \alpha/2$$

$$u = u_{1-\frac{\alpha}{2}} \quad \Rightarrow \quad \frac{\Delta}{\sigma/\sqrt{n}} = u_{1-\frac{\alpha}{2}} \Rightarrow \Delta = \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} \Rightarrow$$

$$P \left\{ \frac{S_n}{n} - \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \theta < \frac{S_n}{n} + \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} \right\} \leq 1 - \alpha$$

Тобто ліва межа довірчого інтервалу

$$\hat{\theta}_1 = \bar{X} - \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

А права межа довірчого інтервалу

$$\hat{\theta}_2 = \bar{X} + \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

Довірчий інтервал з довірчою ймовірністю $1 - \alpha$ для нормально розподіленої генеральної сукупності. Випадок відомої дисперсії

$$\bar{X} - \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \theta < \bar{X} + \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

$u_{1-\frac{\alpha}{2}}$ - квантиль порядку $1 - \frac{\alpha}{2}$; σ^2 - відома дисперсія

$u_{1-\frac{\alpha}{2}}$ - квантиль порядку $1 - \alpha/2$; σ^2 - відома дисперсія

Приклад. Дана вибірка 9, 5, 7, 7, 4, 10, дисперсія $\sigma^2=1$. Побудуємо 99% довірчий інтервал.

Вибіркове середнє дорівнює

$$\hat{\theta} = \bar{X} \equiv \frac{S_n}{n} = \frac{9 + 5 + 7 + 7 + 4 + 10}{6} = 7$$

Довірча ймовірність $0,99 = 1 - \alpha$. Звідси рівень значущості $\alpha = 0,01$. Для побудови довірчого інтервалу потрібен квантиль нормального розподілу порядку $1 - \frac{\alpha}{2} = 1 - \frac{0,01}{2} = 0,995$. За таблицями знаходим

$$u_{1-\frac{\alpha}{2}} = u_{0,995} = 2,58 \text{ .}$$

Обчислюємо

$$\Delta = \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} = \frac{1 \cdot 2,58}{\sqrt{6}} = 1,05$$

Отримаємо довірчий інтервал

$$(7-1,05 ; 7+1,05) = (5,95 ; 8,05)$$

Довірчий інтервал з довірчою ймовірністю $1 - \alpha$ для нормально розподіленої генеральної сукупності. Випадок невідомої дисперсії і малих вибірок ($n < 30$)

$$\bar{X} - \frac{s \cdot t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}} < \theta < \bar{X} + \frac{s \cdot t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}}$$

де $t_{1-\frac{\alpha}{2}}(n-1)$ - квантиль розподілу Стюдента порядку $1 - \frac{\alpha}{2}$ з $n - 1$ ступенями свободи ($df = n - 1$),

$$s^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{(n - 1)}$$

- «підправлена» емпірична дисперсія.

- **Приклад.** Нехай обсяг вибірки $n = 16$, вибіркове середнє $\bar{X} = 5$, «підправлена» вибіркова дисперсія $s^2 = 4$. Побудуємо 99% довірчий інтервал для середнього значення. Оскільки невідома дисперсія генеральної сукупності та $n < 30$, то скористаємося вищенаведеною формулою. Враховуючи, що $\alpha = 1 - 0,99 = 0,01$ за таблицями знайдемо

$$t_{1-\frac{\alpha}{2}}(n-1) = t_{1-\frac{0,01}{2}}(16-1) = t_{0,995}(15) = 3,29$$

Тоді шуканий 99%-довірчий інтервал має вигляд

$$(5-1,645 ; 5+1,645) = (3,355 ; 6,645)$$

Довірчий інтервал з довірчою ймовірністю $1 - \alpha$. Випадок невідомої дисперсії і великих вибірок ($n \geq 30$)

$$\bar{X} - \frac{s \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \theta < \bar{X} + \frac{s \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

$u_{1-\frac{\alpha}{2}}$ - квантиль нормального розподілу порядку $1 - \frac{\alpha}{2}$; s^2 «підправлена» вибіркова дисперсія.

Мінімальний об'єм вибірки для заданої точності.

Як звужити довірчий інтервал.

Часто постає задача збільшити точність (тобто зменшити інтервал) не змінюючи рівень довіри. Цього можна досягти тільки збільшуючи об'єм вибірки. Дійсно, точність залежить від об'єму вибірки

$$\Delta = \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

Найменший об'єм вибірки, що гарантує задану точність (довжину інтервалу):

$$n \geq \left(\frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\Delta} \right)^2$$

Довірчий інтервал з довірчою ймовірністю $1 - \alpha$ для частки ознаки.

Випадок великих вибірок ($n \geq 30$)

Нехай $\hat{p}$ означає частку (відносну частоту) досліджуваної ознаки у вибірці, $\hat{q} = 1 - \hat{p}$. Тоді для частки ознаки в генеральній сукупності θ довірчий інтервал з довірчою ймовірністю $1 - \alpha$ має вид

$$\hat{p} - u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} < \theta < \hat{p} + u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

$u_{1-\frac{\alpha}{2}}$ - квантиль нормального розподілу порядку $1 - \frac{\alpha}{2}$

Для коректного застосування цієї формули потрібно виконання умов

$$n \cdot \hat{p} \geq 5, \quad n \cdot \hat{q} \geq 5$$

Приклад. Обсяг вибірки $n = 100$, вибіркова частка $\hat{p} = 0,2$. Тоді $\hat{q} = 1 - 0,2 = 0,8$. Побудуємо 95% довірчий інтервал для частки ознаки в генеральній сукупності. Рівень значущості

$$\alpha = 1 - 0,95 = 0,05$$

Умови надійності використання цих формул

$$n \cdot \hat{p} = 100 \cdot 0,2 = 20 > 5; \quad n \cdot \hat{q} = 100 \cdot 0,8 = 80 > 5$$

виконуються. За таблицею кватилей нормального розподілу

$$u_{1-\frac{\alpha}{2}} = u_{1-\frac{0,01}{2}} = u_{0,995} = 1.96$$

Тоді шуканий 95%-довірчий інтервал має вигляд

$$(0.2-0.078 ; 0.2+0.078) = (0,122 ; 0.278)$$

Довірчий інтервал з довірчою ймовірністю $1 - \alpha$ для дисперсії

Довірчий інтервал для дисперсії має вигляд

$$\left(\frac{(n-1) \cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} ; \frac{(n-1) \cdot s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right)$$

де s^2 -«підправлена» емпірична дисперсія;

$\chi^2_{\frac{\alpha}{2}}(n-1)$ та $\chi^2_{1-\frac{\alpha}{2}}(n-1)$ -квантілі χ^2 -розподілу з $n-1$ ступенями свободи та відповідно порядку $\frac{\alpha}{2}$ та $1-\frac{\alpha}{2}$.

Приклад. Обсяг вибірки $n = 20$, вибіркова дисперсія $s^2 = 196$.

Побудуємо 90% довірчий інтервал для дисперсії генеральної сукупності (теоретичної дисперсії). Рівень значущості

$$\alpha = 1 - 0,90 = 0,10$$

За таблицею квантилей χ^2 -розподілу

$$\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{0,05}(20-1) = \chi^2_{0,05}(19) = 10,12$$

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{0,95}(20-1) = \chi^2_{0,95}(19) = 30,14$$

Тоді шуканий 90%-довірчий інтервал має вигляд

$$\left(\frac{(20-1) \cdot 196}{30,14}; \frac{(20-1) \cdot 196}{10,12} \right) = (123,56; 366,54)$$

ТЕМА 14. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ В АНАЛІТИЧНИХ ЗАДАЧАХ

ЛЕКЦІЙНЕ ЗАНЯТТЯ № 15

Перевірка гіпотез

У розглянутому прикладі ми стикалися з двома конкуруючими твердженнями. Перше з них - добавка ефективна, а друге – вона не діє. Щоб застосувати статистичний інструментарій треба формалізувати задачу вибору одного з цих тверджень і здійснити такий вибір на основі спостережень.

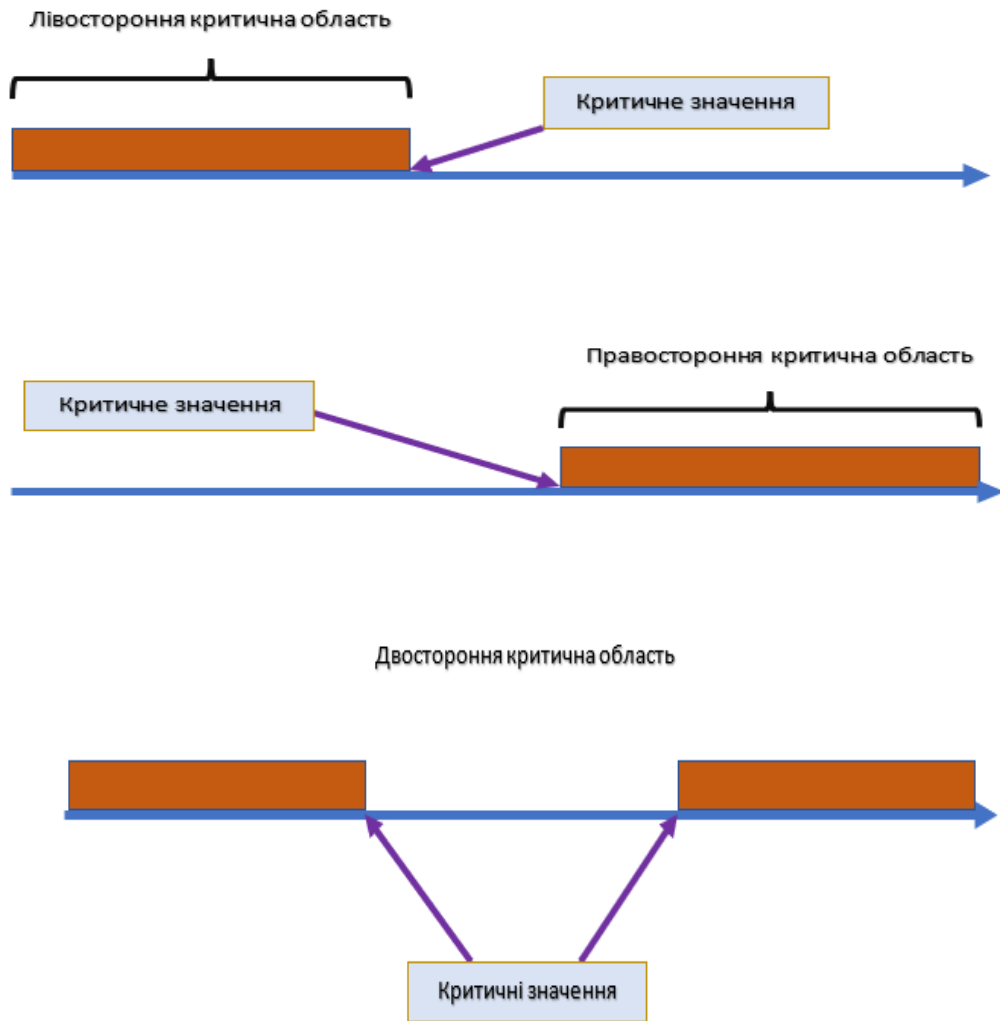
З цією метою розглядаються два твердження, з яких треба вибрати одне на основі деякого правила, яка називається **статистичним критерієм**. Одне з них називають основною або **нульовою гіпотезою** (позначають H_0), а друге називають конкуруючою гіпотезою, або **альтернативною гіпотезою** (позначають H_1).

Числову пряму розбивають на дві частини: область прийняття рішень та критичну область.

Визначення. Область, при попаданні у яку статистики критерію, відкидається нульова (основна) гіпотеза, називається **критичною**.

Доповнення до критичної області (тобто коли приймається нульова гіпотеза) називається **областю прийняття рішень**.

У відповідності до формату критичної області називають критерії, **лівостороннім, правостороннім, двостороннім**.



Становище критичної області залежить від формулювання альтернативної гіпотези. Наприклад, якщо перевіряється гіпотеза

$$H_0: \theta = \theta_0$$

проти альтернативної

$$H_1: \theta > \theta_0$$

то критична область V_1 розміщується праворуч на розподілі статистики Z , і набуває вигляду:

$$Z > z_{1-\alpha}$$

де $z_{1-\alpha}$ – квантиль порядку $1 - \alpha$ розподілу статистики Z за умови, що правильна основна гіпотеза H_0 .

Якщо альтернативна гіпотеза має вигляд

$$H_1: \theta < \theta_0$$

то критична область V_1 розміщується ліворуч на розподілі статистики Z , і набуває вигляду:

$$Z < z_\alpha$$

де z_α – квантиль порядку α розподілу статистики Z за умови, що правильна основна гіпотеза H_0 .

Якщо в гіпотезі сформульовано твердження тільки про одне значення параметру (або функції розподілу), то така гіпотеза називається простою. В іншому випадку гіпотеза називається складною.

Приклад 1. Для рішення про доцільність виходу на ринок з новим товаром керівництво фірми вважає, що частка споживачів, що висловлюють бажання купити продукту в генеральній сукупності не повинна бути менше 20%. Це припущення, яке потрібно перевірити, сформулюється у вигляді гіпотези про параметр розподілу p дихотомічно розподіленої генеральної сукупності:

$$H_0: p \geq 0.2$$

Це приклад складної гіпотези. Як альтернативну гіпотезу можна вибрати одну з гіпотез:

$$H_1: p = 0.1$$

$$H_1: p < 0.2$$

Перша з цих гіпотез проста, а друга - складна.

Оскільки статистика критерію є випадковою величиною, то рішення про прийняття або відхилення основної гіпотези може бути помилковим.

Помилка 1-ого роду: помилково відхиляється основна гіпотеза (тобто приймається альтернативна гіпотеза, але вірна нульова гіпотеза). Ймовірність помилки першого роду позначають α і називають рівнем значущості:

$$P(Z \in V_1 / H_0) = \alpha$$

Приклад 2. Відомо, що генеральна сукупність має нормальний розподіл $N(\theta, 1)$, де $\theta = EX$ – невідоме математичне сподівання, щодо якого сформульовані гіпотези; $1 = DX$ – відома дисперсія. Задаємо рівень значущості $\alpha=0,05$.

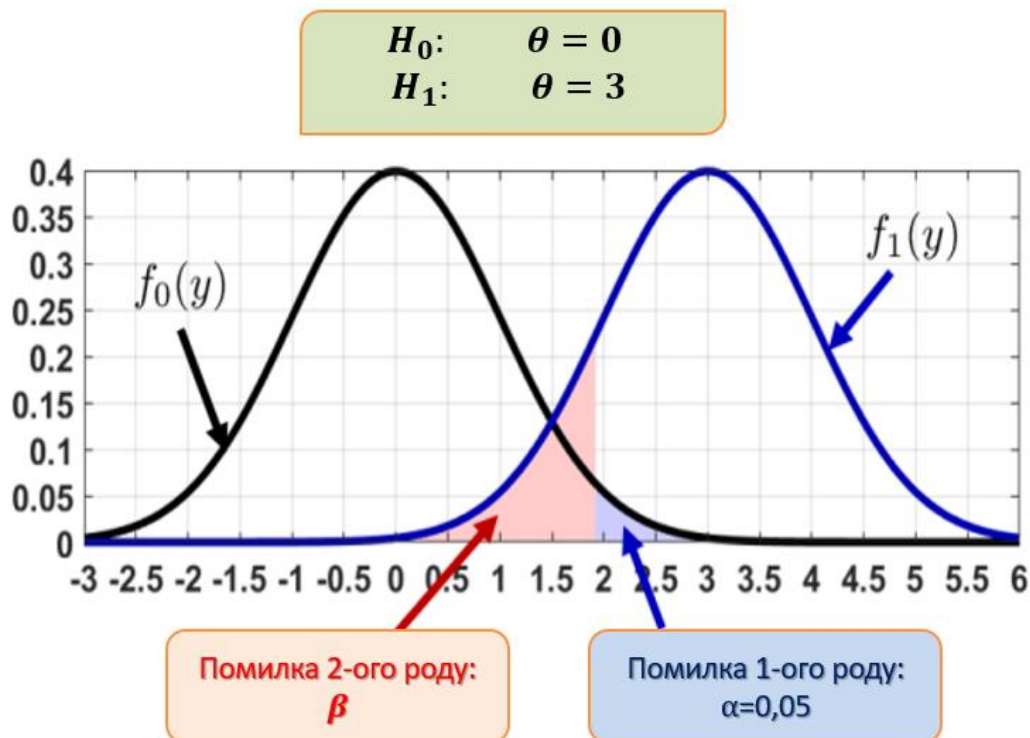
Перевіряється нульова гіпотеза

$$H_0: \quad \theta = 0$$

проти альтернативної гіпотези

$$H_1: \quad \theta = 3$$

Візуалізуємо помилки першого та другого роду.



Джерело: Chan S.H. *Introduction to Probability for Data Science*. 2021. Michigan Publishing Services. 704 p.

Величина $1 - \beta$ називається потужністю критерія.



Джерело: Chan S.H. *Introduction to Probability for Data Science*. 2021. Michigan Publishing Services. 704 p.

Логічна схема статистичного критерію (алгоритм перевірки гіпотез)

1. Висувається нульова гіпотеза H_0 та альтернативна H_1 .
2. Задається рівень значущості α .
3. Обирається критерій (статистика критерія) і визначається розподіл статистики критерію.
4. Будується критична область за допомогою критичних значень (які знаходяться з таблиць).
5. Обчислюється статистика критерію.
6. Приймається рішення про вибір або відхилення гіпотез: якщо статистика критерію попадає в критичну область, то приймається альтернативна гіпотеза; якщо статистика критерію попадає в область прийняття рішень, то відхиляється альтернативна гіпотеза і приймається нульова.

Stochastic Foundations of Data Analysis. Lecture Materials [Electronic resource] : study guide for bachelor's degree students in the educational program "Industrial Marketing", specialty D5 Marketing / Igor Sikorsky Kyiv Polytechnic Institute ; by S. Solntsev, O. Chernenko, O. Kofanov – Electronic text data (1 file) – Kyiv : Igor Sikorsky Kyiv Polytechnic Institute, 2025. 160 p.

ABSTRACT

The contemporary information landscape is characterized by a rapid increase in data volumes and the growing role of intelligent data analysis in solving scientific, economic, and managerial tasks. Effective decision-making in the real sector of the economy, finance, marketing, and business increasingly relies on data-driven strategies, forecasting, and statistical risk assessment. One of the key tools for such analysis is stochastic methods, which allow building probabilistic models, assessing uncertainty, and identifying hidden patterns in large datasets. The application of these methods serves as the foundation for modern analytical tools and machine learning algorithms, contributing to improved accuracy in economic forecasts and the development of optimal management strategies.

In many fields, particularly in industrial marketing and economics, the need for stochastic analysis is driven by the multidimensionality and complexity of real-world processes that generate vast data streams with numerous random factors. Forecasting fluctuations in market indicators, modeling supply and demand dynamics, conducting in-depth consumer behavior research, and analyzing financial risks – all require a deep understanding of probabilistic patterns and sound parameter estimation methods. Knowledge of the fundamentals of stochastic modeling enables the analysis of such processes based on quantitative metrics, which is an essential

component of training modern specialists in marketing, economics and business analytics.

This textbook is designed to provide higher education students with fundamental knowledge of the stochastic foundations of data analysis and to develop practical skills in working with economic datasets. It covers the core sections necessary for a deep understanding of applied modeling, namely basic stochastic event models, the application of combinatorics, Bayesian approaches for updating forecasts, stochastic modeling of discrete and continuous features, numerical characteristics of data distributions, as well as methods of data summarization, statistical estimation, and hypothesis testing. Special attention is given to the algorithms and statistical methods used in real-world economic research, allowing students to acquire the competencies required to apply the gained knowledge in their professional careers.

The purpose of this textbook is to equip students with fundamental knowledge and practical skills in stochastic analysis for effectively solving analytical problems related to process research, large-scale data processing, and informed decision-making under conditions of market uncertainty. In turn, the subject encompasses stochastic processes and patterns that emerge from the mass interaction of random factors in a business environment, along with mathematical algorithms for systematization, hypothesis testing, and the utilization of statistical data to draw practical conclusions.

Mastering the material in this textbook will contribute to the formation of essential competencies, including the capacity for abstract and analytical thinking, the ability to utilize mathematical frameworks for developing predictive models, and the enhancement of critical data literacy skills. Understanding the principles of stochastic modeling allows one not only to assess the degree of uncertainty in decision-making processes but also to take

a reasoned approach to selecting optimal strategies in a highly competitive digital environment.

The textbook is aligned with the program learning outcomes formulated in the course syllabus, serving as a vital theoretical foundation for further study in disciplines dedicated to econometrics, statistics, business analytics, and artificial intelligence systems. The manual contains structured methodological information, supplemented by illustrative examples and algorithmic flowcharts, which facilitate a better understanding of the mathematical tools and their application in real-world scenarios.

© Sergii Solntsev, Oksana Chernenko, Oleksii Kofanov, 2025

National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

ВИКОРИСТАНІ І РЕКОМЕНДОВАНІ ЛІТЕРАТУРНІ ДЖЕРЕЛА

1. Веригіна І. В., Островська О. В., Сугакова О. В. (2022) *Теорія ймовірностей та математична статистика. Лекції і практикум*. Київ : КПІ ім. Ігоря Сікорського, 254 с. URL: <https://ela.kpi.ua/handle/123456789/51552>
2. Гіхман І.І., Скороход А.В., Ядренко М.Й. (1988) *Теорія ймовірностей та математична статистика*. Київ: «Вища школа». 439 с.
3. Горбачук В. М., Кушлик-Дивульська О. І. (2022) *Теорія ймовірностей та математична статистика : підручник для здобувачів ступеня бакалавра за технічними та економічними спеціальностями*. Київ : КПІ ім. Ігоря Сікорського, 351 с. URL: <https://ela.kpi.ua/handle/123456789/52357>
4. Клесов О.І. (2010) *Вибрані питання теорії ймовірностей та математичної статистики: навчальний посібник*. К.: ТВіМС. 248 с.
5. Кофанов О. Є., Солнцев С. О., Зозульов, О. В. (2023) *Програмування із використанням R у статистичних та маркетингових дослідженнях: Навчально-методичний комплекс дисципліни [Електронний ресурс] : навч. посіб. для студентів спеціальності 075 «Маркетинг» КПІ ім. Ігоря Сікорського. Електронні текстові дані (1 файл: 4.74 Мбайт)*. Київ : КПІ ім. Ігоря Сікорського, 204 с. URL: <https://ela.kpi.ua/handle/123456789/56233>
6. Куц Ю. В., Лисенко Ю. Ю. (2022) *Спеціальні розділи математики. Курс лекцій : навчальний посібник для здобувачів ступеня бакалавра*. Київ: КПІ ім. Ігоря Сікорського, 2022. 180 с. URL: <https://ela.kpi.ua/handle/123456789/48823>
7. Кушлик-Дивульська О. І., Защепкіна Н. М. (2022) *Спецрозділи вищої математики : навчальний посібник*. Київ : КПІ ім. Ігоря Сікорського, 2022. 148 с. URL: <https://ela.kpi.ua/handle/123456789/50171>
8. Скороход А.В. (1980). *Ймовірність навколо нас*. Київ: «Наукова думка», 196 с.

9. Солнцев С. О., Жигалкевич Ж. М. (2020) *Визначення привабливості інноваційних проектів в межах квазіструктур на основі коджойнт підходу. Маркетинг і цифрові технології*. 2020. 1. С. 15-28. URL: <https://doi.org/10.30525/2256-0742/2020-6-3-124-135>
10. Солнцев С. О., Москаленко О. Д., Черненко О. В. (2018) *Система моніторингу маркетингового середовища підприємства. Економічний вісник НТУУ «КПІ»*. 2018. № 15. С. 341–354.
11. Солнцев С. О., Черненко О. В. (2020) *Теорія ймовірностей та математична статистика: навчально-методичний комплекс*. Київ : КПІ ім. Сікорського, 2020. 137 с. URL: <https://ela.kpi.ua/handle/123456789/43424>
12. Солнцев С.О. , Жигалкевич Ж.М. (2023) *Моделювання управління економічними системами і процесами. Навчально-методичний комплекс дисципліни (перероблений і доповнений) [Електронний ресурс] : навч. посіб. для здобувачів ступеня доктора філософії за освітньо-науковою програмою «Менеджмент» спеціальності 073 «Менеджмент» / КПІ ім. Ігоря Сікорського ; уклад.: С. О. Солнцев, Ж. М. Жигалкевич. Електронні текстові дані (1 файл: 1.14 Мбайт). Київ : КПІ ім. Ігоря Сікорського, 75 с. URL: <https://ela.kpi.ua/handle/123456789/57509>*
13. Buldygin V.V., Solntsev S.A. (1987) *The contraction principle and the strong law of large numbers for weighted sums*. Theory of Probability & Its Applications. 31 (3), 452-462 p. (Web of Science)
14. Buldygin V.V., Solntsev S.A. (1988) *The strong law of large numbers for sums of independent random vectors with operator normalizations*. Theory of Probability & Its Applications. 32 (2), 243-256. (Web of Science)
15. Buldygin V.V., Solntsev S.A. (1997) *Asymptotic Behaviour of Linearly Transformed Sums of Random Variables*. Kluwer Academic Publishers, 500 p.
16. Chan S.H. (2021) *Introduction to Probability for Data Science*. Michigan Publishing Services, 704 pp.
17. Mosteller F. (1987) *Fifty Challenging Problems in Probability with Solutions*. Dover Publications. 96 p.
18. Siegel, A.F. (2016) *Practical Business Statistics*. Elsevier, 622 p.

- 19. Solntsev S., Zhygalkevych Zh., Kravchenko M.** *Evaluation of risk impact on implementation of innovation projects within the framework of machine-building quasi-integration structures. Baltic Journal of Economic Studies.* 2020. 6(3). P. 124-135. URL: <https://doi.org/10.30525/2256-0742/2020-6-3-124-135>
- 20. Székely G.J.** (1986) *Paradoxes in Probability Theory and Mathematical Statistics.* Reidel, Kluwer Academic Publishers, 264 p.
- 21. Zgurovsky M., Gavrysh O., Solntsev S., Kukharuk A., Skorobogatova N.** *Selection of indicators for scenario modeling of countries economy development. Problems and perspectives in management.* 2020. Vol. 3. P. 1810-5467. URL: [http://dx.doi.org/10.21511/ppm.18\(2\).2020.36](http://dx.doi.org/10.21511/ppm.18(2).2020.36)

ДОДАТКИ

ДОДАТОК А

Таблиця А.1 Розподіл Стюдента $t_{\alpha,n}$ для різних значень довірчої імовірності p
та числа вимірів n

n	0,5	0,6	0,7	0,8	0,9	0,95	0,98	0,99
2	1.000	1,376	1,963	3,08	6,31	12,71	31,8	63,7
3	0,816	1,061	1,336	1,886	2,92	4,30	6,96	9,92
4	0,765	0,978	1,250	1,638	2,35	3,18	4,54	5,84
5	0,741	0,941	1,190	1,533	2,13	2,77	3,75	4,60
6	0,727	0,920	1,156	1,476	2,02	2,57	3,36	4,03
7	0,718	0,906	1,134	1,440	1,943	2,45	3,14	4,71
8	0,711	0,896	1,119	1,415	1,895	2,36	3,00	3,50
9	0,706	0,889	1,108	1,397	1,860	2,31	2,90	3,36
10	0,703	0,883	1,110	1,383	1,833	2,26	2,82	3,25
11	0,700	0,879	1,093	1,372	1,812	2,23	2,76	3,17
12	0,697	0,876	1,088	1,363	1,796	2,20	2,72	3,11
13	0,695	0,873	1,083	1,356	1,782	2,18	2,68	3,06
14	0,694	0,870	1,079	1,350	1,771	2,16	2,65	3,01
15	0,692	0,868	1,076	1,345	1,761	2,14	2,62	2,98
16	0,691	0,866	1,074	1,341	1,753	2,13	2,60	2,95
17	0,690	0,865	1,071	1,337	1,746	2,12	2,58	2,92
18	0,689	0,863	1,069	1,333	1,740	2,11	2,57	2,90
19	0,688	0,862	1,067	1,330	1,734	2,10	2,55	2,88
20	0,688	0,861	1,066	1,328	1,729	2,09	2,54	2,86
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,33	2,58

Критичні значення F-критерію (при $\alpha = 0,05$)

Ступені свободи, k_1	Ступені свободи, k_2						
	1	2	3	4	5	6	50
9	5,12	4,26	3,86	3,63	3,48	3,37	2,8
10	4,96	4,10	3,71	3,48	3,33	3,22	-
11	4,84	3,98	3,59	3,36	3,20	3,09	2,5
12	4,75	3,88	3,49	3,26	3,11	3,00	-
13	4,67	3,80	3,41	3,18	3,02	2,92	2,32
14	4,60	3,74	3,34	3,11	2,96	2,85	-
15	4,54	3,68	3,29	3,06	2,90	2,79	2,18
16	4,49	3,63	3,24	3,01	2,85	2,47	-
17	4,45	3,59	3,20	2,96	2,81	2,70	2,08
18	4,41	3,56	3,16	2,93	2,77	2,66	-
19	4,33	3,52	3,13	2,90	2,74	2,63	2,08
20	4,35	3,49	3,10	2,87	2,71	2,90	-
21	4,32	3,47	3,07	2,84	2,68	2,57	1,93
22	4,30	3,44	3,05	2,82	2,66	2,55	-
23	4,28	3,42	3,03	2,80	2,64	2,53	1,88
24	4,26	3,40	3,01	2,78	2,62	2,51	-
25	4,24	3,38	2,99	2,76	2,60	2,49	1,84
26	4,22	3,37	2,98	2,74	2,59	2,47	-
27	4,21	3,35	2,96	2,73	2,57	2,46	1,80
28	4,20	3,34	2,95	2,71	2,56	2,44	-
29	4,18	3,38	2,93	2,70	2,54	2,43	1,77
30	4,17	3,32	2,92	2,69	2,53	2,42	-
35	4,11	3,26	2,86	2,63	2,48	2,36	-
40	4,08	3,23	2,84	2,61	2,45	2,34	-
50	4,03	3,18	2,79	2,56	2,40	2,29	-
60	4,00	3,15	2,76	2,52	2,37	2,25	-
∞	3,84	2,99	2,60	2,37	2,21	2,09	-

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

ДЛЯ НОТАТОК

[illegible]

ДЛЯ НОТАТОК

[illegible]