ТЕОРІЯ ЙМОВІРНОСТЕЙ

СКАСКІВ ОЛЕГ

ЛЕКЦІЇ 1 – 14

14.05.2021

ЛЕКЦІЯ 1. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ

1.1. Випадкове явище (випадковий експеримент)

У подальшому ми розрізнятимемо експерименти (чи явища) детерміновані та стохастичні= випадкові.
Детермінованими, експериментами (чи явищами) називатимемо експерименти (чи явища), для яких при відтворенні одного і того ж комплексу умов, що забезпечують проходження експерименту, достовірно (точно) можна передбачити результат експерименту.
Стохастичним або випадковим експериментом називають експерименти (явища), перебіг яких і їхні результати не піддаються точному прогнозуванню. Такі експерименти або явища, які називають стохастичними (випадковими) експериментами або явищами, і є предметом розгляду теорії ймовірностей.

До стохастичних явищ (експериментів) на початковому етапі моделювання їхнього перебігу відносять також недостатньо вивчені явища, внаслідок чого у нас немає всієї повноти інформації про характеристики, що описують їх перебіг. До явищ (експериментів) такого ж типу відносяться явища (експерименти), всю інформацію про перебіг яких з різних мотивів неможливо (або й недоцільно) збирати.
Отже, початковими поняттями у нас виступають поняття випадкового явища (експерименту), тобто явища (експерименту), результат перебігу якого при відтворенні одного і того ж комплексу умов, що забезпечують його "проведення", неможливо точно передбачити.
З кожним експериментом (явищем) пов'язуємо простір Ω всіх, які тільки собі можна уявити, результатів експерименту.
Точки $\omega \in \Omega$ цього простору називаються елементарними подіями, сам простір називають простір елементарних подій.

Розглянемо спочатку приклади просторів елементарних подій для кількох модельних експериментів. Важливою обставиною при цьому є точне описання умов (обставин), що забезпечують проведення (відтворення) експерименту.

Підкидання монети

1. Так, для експерименту з підкидання симетричної плоскої (без товщини) монети на ідеально рівну горизонтальну поверхню простір елементарних подій $\Omega = \{Г, Ц\}$, тобто, Г — "падіння монети гербом догори", Ц — "падіння монети цифрою догори".

2. Разом з цим підкидання монети номіналом 1 гривня чи 1 фунт (стерлінгів) дає, як нескладно перевірити, ще одну можливість "падіння на ребро" — Р, і простір елементарних подій у цьому випадку має три точки $\Omega = \{Г, Ц, Р\}$. Зазначимо, що ми не пояснили, що ж ми розуміємо, коли говоримо "підкидання", а поклалися на інтуїтивне розуміння того, що це підкидання повинно здійснюватися довільно, без контролю з боку експериментатора.

Підкидання грального кубика

При підкиданні симетричного грального кубика на ідеально рівну горизонтальну поверхню простір елементарних подій можна описати у вигляді

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

де цифри означають кількість очок на грані кубика, яка виявилась згори. Можливий, наприклад, такий результат цього експерименту: A — "випадання парної кількості очок". Цю подію можна описати, як таку множину

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

Підкидання (повторне) монети і грального кубика

Якщо підкидаємо двічі симетричну монету, то простір елементарних подій описуємо як множину пар $\Omega = \{\text{ГГ}, \text{ЦЦ}, \text{ЦГ}, \text{ГЦ}\}$, а якщо гральний кубик підкидаємо двічі, то отримуємо простір елементарних подій

$$\Omega = \{(i, j) : 1 \leq i \leq 6,\ 1 \leq j \leq 6\},$$

тобто, як множину впорядкованих пар кількості очок, що виявились на верхній грані кубика першого і другого разу, відповідно.

Якщо підкидаємо тричі симетричну монету, то простір елементарних подій описуємо як множину

$$\Omega = \{\text{ГГГ}, \text{ЦЦЦ}, \text{ЦГГ}, \text{ГГЦ}, \text{ГЦГ}, \text{ЦГЦ}, \text{ЦЦГ}, \text{ГЦЦ}\},$$

а якщо гральний кубик підкидаємо тричі, то отримуємо простір елементарних подій

$$\Omega = \{(i, j, k) : 1 \leq i, j, k \leq 6\}.$$

Стрільба по круглій мішені.

У цьому експерименті (Стрільба по круглій мішені) за простір елементарних подій можна вибрати

$$\Omega = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq R^2\},$$

де R — радіус мішені. Цей експеримент у загальнішій формі можна ще описати як вибір “навмання точки” з "тіла" Ω в Евклідовому просторі $\mathbb{R}^d$ $(d \geq 1)$.

У перших чотирьох описаних прикладах природно будь-яку підмножину $A \subset \Omega$ називати подією. У останніх двох прикладах подія — не кожна підмножина. Подіями тут можна вважати лише у деякому сенсі вимірні підмножини Ω (наприклад, борелеві). В загальному ж “подіями” в даному експерименті, якому відповідає простір елементарних подій Ω, природно вважати такі підмножини $A \subset \Omega$, для яких за умов, що забезпечують перебіг експерименту, є можлива лише одна з відповідей такого типу: $\omega \in A$ або $\omega \notin A$. Зрозуміло, що тоді Ω, $\varnothing$ — події у даному експерименті, перша з яких Ω — ”достовірна подія”, тобто така подія, що відбувається при кожному відтворенні експерименту, і подія, яка визначається порожньою множиною $\varnothing$ — ”неможлива подія”, тобто така

1.2. Операції над подіями

Отже, попередній і дуже початковий інтуїтивний аналіз випадкових явищ привів нас до розуміння того, що весь простір Ω і порожню множину $\varnothing$ обов'язково слід внести у перелік підмножин Ω, які ми розглядатимемо, як події. Водночас, якщо підмножина $A \subset \Omega$ є подією, то подію повинна визначати і множина $\overline{A} \stackrel{def}{=} \{\omega \in \Omega \colon \omega \notin A\} = \Omega \setminus A$ — протилежна подія до A, тобто, $\overline{A}$ це подія, яка полягає в тому, що подія A не відбувається. Наступні операції над подіями також є елементарними вимогами до сім'ї множин, лише які ми розглядатимемо як події, що випливають із загальних властивостей класів множин.

$A + B \stackrel{def}{=} A \cup B$ — сума подій, тобто, подія, яка полягає у тому, що відбувається подія A або подія B.

$A \cdot B = A \cap B$ — добуток подій, тобто, подія, яка полягає в тому, що одночасно відбуваються події A та B.

$A \setminus B$ — різниця подій, тобто, подія яка полягає у тому, що перша подія відбувається і друга подія не відбувається.

Ще крім цього, події A і B називатимемо несумісними подіями, якщо $A \cdot B = A \cap B = \varnothing$ (тобто якщо добуток цих двох подій — неможлива подія).

Повернемось до того, як можна підійти до оцінювання шансів (ймовірностей) появи тих або інших подій у даному експерименті. Нехай A — деяка подія, за появою якої проводиться спостереження у експерименті. Шансові появи події хочемо співставляти число з проміжку $[0, 1]$ так, що шанс появи достовірної події дорівнює одиниці, а шанс появи неможливої події дорівнює нулю. Якщо $n \in \mathbb{N}$ — кількість повторень даного експерименту, а $m_n(A)$ — кількість появ події A, то природною характеристикою події A є частота її появи

$$\nu_n(A) = \frac{m_n(A)}{n} \text{ — частота появи події } A.$$

Гіпотетично можна припустити, що є події для яких існує границя

$$\lim_{n\to\infty} \nu_n(A) = \mathbb{P}(A).$$

Тоді цю границю природно оголосити ймовірністю події A. Це класичний підхід до означення ймовірності. Однак, при цьому виникає проблема опису можливих результатів експерименту, для яких така границя існує. Ми не будемо розвивати далі цей, історично перший підхід, до побудови моделей випадкових явищ. У даний момент вважається природним вводити поняття ймовірності дещо інакше. Перш, ніж перейдемо до побудови системи аксіом теорії ймовірностей, відзначимо деякі властивості ймовірності $\mathbb{P}(A)$, які випливають з даного щойно "статистичного" означення і, які бажано, щоб ймовірність початково задовольняла при її аксіоматичному запровадженні. Власне, спробуємо з'ясувати, чому ймовірність $\mathbb{P}(A)$ повинна бути мірою?

Відзначимо такі властивості $\mathbb{P}(A)$:

1. $0 \leq \nu_n(A) \leq 1 \Rightarrow 0 \leq \mathbb{P}(A) \leq 1$.
2. $\nu_n(\Omega) = 1 \Rightarrow \mathbb{P}(\Omega) = 1$. Отже, ймовірність достовірної події дорівнює 1.
3. $A \cdot B = \varnothing \Rightarrow m_n(A + B) = m_n(A) + m_n(B) \Rightarrow$ $\nu_n(A + B) = \nu_n(A) + \nu_n(B) \Rightarrow \mathbb{P}(A + B) = \lim\limits_{n \to +\infty} \nu_n(A + B) =$ $= \lim\limits_{n \to +\infty} \nu_n(A) + \lim\limits_{n \to +\infty} \nu_n(B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Висновок, який напрошується з щойно наведених простих імплікацій такий — ймовірність $\mathbb{P}(A)$ при аксіоматичному її визначенні повинна бути адитивною функцією (адитивною мірою) події з ймовірністю всього простору, що дорівнює 1, тобто, так званою, ймовірнісною мірою.

Власне, ми з'ясували, що мінімальні умови, які повинна задовольняти сім'я подій зводиться до виконання аксіом алгебри сім'ї підмножин простору елементарних подій, а ймовірність, як числова функція події, яка набуває значення з проміжку $[0, 1]$, повинна бути ймовірнісною адитивною мірою.

АКСІОМАТИЧНЕ ВИЗНАЧЕННЯ ЙМОВІРНОСТІ

Ми розглядали такі моделі, в яких множина елементарних виходів (результатів) експерименту скінченна, а також, без достатнього на те обґрунтування, — геометричні ймовірності. Якщо в першому випадку будь-яка сукупність результатів експерименту трактувалась як подія, то у випадку геометричних ймовірностей ми були вимушені вважати подіями лише вимірні за Лебегом множини.

В основу всього подальшого викладу покладено поняття ймовірнісного простору $(\Omega, \mathcal{A}, \mathbb{P})$, де Ω — простір елементарних подій,

$\mathcal{A}$ — σ-алгебра підмножин з Ω, $\mathbb{P}$ — ймовірність (ймовірнісна зліченно-адитивна міра).

Нехай простір елементарних подій (елементарних результатів експерименту) Ω є довільною множиною, 2^{Ω} — множина всіх підмножин множини Ω, а $\mathcal{A} \subset 2^{\Omega}$ — деяка система підмножин множини Ω. Для $A \subset \Omega$ позначатимемо $\overline{A} = \Omega \setminus A$.

Означення (алгебри підмножин)

Сім'я $\mathcal{A} \subset 2^{\Omega}$ називається алгеброю, якщо для її елементів виконуються такі аксіоми: 1. $\Omega \in \mathcal{A}$; 2. $(\forall A)\colon A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$;
3. $(\forall A, B)\colon A, B \in \mathcal{A} \Rightarrow A + B \in \mathcal{A}$.

Означення (σ-алгебри підмножин)

Сім'я $\mathcal{A} \subset 2^{\Omega}$ називається σ-алгеброю, якщо виконуються такі умови:

1. $\Omega \in \mathcal{A}$; 2. $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$; 3. $(\forall (A_j)) : A_j \in \mathcal{A}\ (j \geq 1) \Rightarrow \bigcup\limits_{j=1}^{+\infty} A_j \in \mathcal{A}$.

Означення (скінченно-адитивної ймовірнісної міри)

Нехай $\mathcal{A}_0$ — алгебра. Відображення $\mathbb{P}_0\colon \mathcal{A}_0 \to \mathbb{R}$ називається скінченно-адитивною ймовірнісною мірою на $\mathcal{A}_0$, якщо виконуються такі умови: 1. $(\forall A \in \mathcal{A}_0) : \mathbb{P}_0(A) \geq 0$; 2.

$(\forall n \in \mathbb{N})(\forall (A_j)_{j=1}^{n}, A_j \in \mathcal{A}_0,\ A_j \cap A_k = \varnothing\ (j \neq k))\colon\ \mathbb{P}_0(\bigcup\limits_{j=1}^{n} A_j) = \sum\limits_{j=1}^{n} \mathbb{P}_0(A_j)$;

3. $\mathbb{P}(\Omega) = 1$.

Означення (зліченно-адитивної міри)

Нехай $\mathcal{A}_0$ — алгебра. Відображення $\mathbb{P}_0 \colon \mathcal{A}_0 \to \mathbb{R}$ називається зліченно-адитивною мірою на $\mathcal{A}_0$, якщо виконуються такі умови:

1. $(\forall A \in \mathcal{A}_0) : \mathbb{P}_0(A) \geq 0$;
2. $(\forall (A_j), A_j \in \mathcal{A}_0,\ A_j \cap A_k = \varnothing\ (j \neq k))$:

$$\bigcup_{j=1}^{+\infty} A_j \in \mathcal{A}_0 \Rightarrow \mathbb{P}_0(\bigcup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \mathbb{P}_0(A_j).$$

Означення

Нехай $\mathcal{A}_0$ — алгебра. Відображення $\mathbb{P} \colon \mathcal{A}_0 \to \mathbb{R}$ називається зліченно-адитивною ймовірнісною мірою на алгебрі $\mathcal{A}_0$, якщо:

1. $\mathbb{P}$ — зліченно-адитивна міра на $\mathcal{A}_0$; 2. $\mathbb{P}(\Omega) = 1$.

Зліченно-адитивну ймовірнісну міру $\mathbb{P}$, визначену на σ-алгебрі називатимемо ймовірністю.

Нехай Ω — простір елементарних подій, $\mathcal{A}$ — σ-алгебра підмножин Ω, $\mathbb{P}$ — ймовірність, визначена на $\mathcal{A}$. Пару $(\Omega, \mathcal{A})$ називатимемо ймовірнісною моделлю експерименту.

Трійка $(\Omega, \mathcal{A}, \mathbb{P})$ називається ймовірнісним простором.

В багатьох задачах досить, щоб $\mathcal{A}$ була алгеброю.

Деякі ймовірнісні моделі

1. Ймовірнісна модель експерименту з не більш, ніж зліченною кількістю подій чи результатів.
У цьому випадку $\Omega = \{\omega_j : 1 \le j \le N\}$, $N \le +\infty$, $\mathcal{A} = 2^{\Omega}$ — σ-алгебра.
Тоді $\forall A \in \mathcal{A} : A = \cup_{\omega_j \in A}\{\omega_j\}$ і, тому, $\mathbb{P}(A) = \sum_{\omega_j \in A} \mathbb{P}(\{\omega_j\})$.
Зрозуміло, що для того, щоб обчислити ймовірність довільної події, достатньо знати лише ймовірності елементарних подій $0 \le p_j \stackrel{def}{=} \mathbb{P}(\{\omega_j\})$, при цьому, оскільки $\Omega = \bigcup_{j=1}^{N} \omega_j$ і $\mathbb{P}(\Omega) = 1$, то повинна виконуватись рівність $\sum_{j=1}^{N} p_j = 1$.

2. Класичні ймовірності. Класичну модель отримаємо, коли $N < +\infty$ і $p_j = p$ $(1 \le j \le N)$, тобто, всі елементарні події є однаково можливими. У цьому випадку $1 = \sum_{j=1}^{N} p_j = p \cdot N$. Тому, $p = \frac{1}{N}$. Звідси, для знаходження ймовірності довільної події $A \in 2^{\Omega}$ отримуємо формулу класичних ймовірностей

$$\mathbb{P}(A) = \sum_{\omega_j \in A} p_j = \sum_{\omega_j \in A} \frac{1}{N} = \frac{|A|}{|\Omega|} = \frac{m}{N},$$

де m — кількість елементарних подій, сприятливих для A, тобто

$$A = \{\omega_{j_1}, \omega_{j_2}, \ldots, \omega_{j_m}\}, \quad \omega_{j_s} \ne \omega_{j_k} \ (s \ne k).$$

Приклад. Двічі підкидаємо гральний кубик. Обчислимо ймовірність того, що при цьому випаде парна сума очок.
$\Omega = \{(i,j) : 1 \le i,j \le 6\}$, $A = \{(i,j) \colon i+j \in \{2,4,6,8,10,12\}\}$ — подія, ймовірність якої потрібно знайти. За формулою класичних ймовірностей

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{18}{36} = \frac{1}{2}.$$

Приклад. Монета підкидається до першої появи герба. Опишемо ймовірнісний простір. Простір елементарних подій

$$\Omega = \{\text{Г}, \text{ЦГ}, \text{ЦЦГ}, \ldots, \omega_j, \ldots\}, \;\; \omega_j = \underbrace{\text{ЦЦЦ}\ldots\text{Ц}}_{j-1}\text{Г},$$

Для знаходження ймовірності $p_j = \mathbb{P}(\omega_j)$ $(j \ge 2)$ розглянемо простір елементарних однаково можливих подій

$$\Omega_j = \{\underbrace{\text{Ц}\ldots\text{Ц}}_{j}, \underbrace{\text{Ц}\ldots\text{Ц}}_{j-1}\text{Г}, \;\ldots\;, \underbrace{\text{Г}\ldots\text{Г}}_{j}\}.$$

Тоді $|\Omega_j| = 2^j$ і за формулою класичних ймовірностей $p_j = \mathbb{P}(\omega_j) = 2^{-j}$ $(j \ge 2)$. Зрозуміло, що $\sum_{j=1}^{\infty} p_j = 1$.

3. Геометричні ймовірності.

Нехай Ω — обмежена вимірна за Лебегом множина в $\mathbb{R}^p$, $p \geq 1$, $\mathcal{A}$ — σ-алгебра вимірних за Лебегом підмножин Ω, зокрема, лише борелевих $\mathcal{B}(\mathbb{R})$. Тоді для $A \in \mathcal{A}$: $\mathbb{P}(A) \stackrel{def}{=} \dfrac{\mathrm{meas}_p(A)}{\mathrm{meas}_p(\Omega)}$,

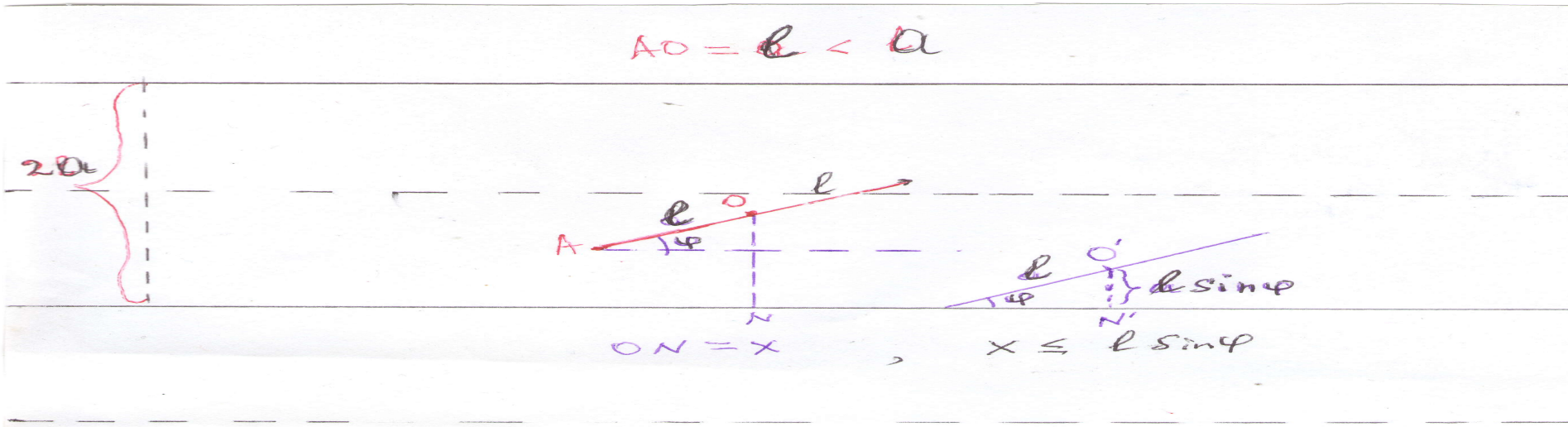
де meas_p — лебегова міра в $\mathbb{R}^p$. Неважко переконатись, що $\mathbb{P}$ — справді ймовірність на $\mathcal{A}$.

Приклад (стрільба по круговій мішені).

$\mathbb{P}(\{10\}) = \dfrac{\pi r^2}{\pi R^2} = \left(\dfrac{r}{R}\right)^2$ — ймовірність попадання в "десятку де r, R — радіуси "десятки" і мішені, відповідно.

Задача Бюффона. На площину, розліновану паралельними прямими, відстань між двома сусідніми з яких дорівнює $2a$ навмання кидаємо голку довжиною 2ℓ, $\ell < a$. Знайти ймовірність того, що голка перетне одну з паралельних ліній.

Нехай φ — кут, що утворює голка з деяким фіксованим напрямком на прямих, $0 \leq \varphi \leq \pi$, x — відстань від центру голки до найближчої з паралельних прямих, $0 \leq x \leq a$. Пара чисел (φ, x) повністю визначає положення голки між двома сусідніми прямими стосовно найближчої прямої.



тарних подій). Опишемо множину тих пар (φ, x), які відповідають події A

— голка перетинає одну з прямих. Очевидно, що лише, коли $x \leq \ell \sin\varphi$ — голка перетинає одну з прямих. Тому,
$A = \{(\varphi, x) : 0 \leq x \leq \ell \sin\varphi, 0 \leq \varphi \leq \pi\}$. Шукана ймовірність того, що голка перетинає лінію

$$\mathbb{P}(A) = \frac{S(A)}{\pi a} = \frac{\ell}{\pi a}\int_0^{\pi} \sin\varphi d\varphi = \frac{2\ell}{\pi a}.$$

Неважко помітити, що при складанні моделі даного експерименту ми зробили декілька припущень, що спростили наші міркування. Зокрема, при побудові простору елементарних подій ми ніяк не обгрунтували переходу від розгляду розлінованої паралельними прямими площини до двох сусідніх паралельних прямих, а також того, що у нашому розгляді ми не розрізняємо цих прямих. З останньою обставиною все доволі просто вирішується, позаяк у наших міркуваннях ці дві прямі є цілком рівноправними, то розгляд відстані "до найближчої з паралельних прямих" замість пари відстаней до прямих, очевидно, що ніяк не змінює числове значення ймовірності (Вправа. Обгрунтувати це твердження!). З першою ж обставиною все значно складніше.

Чи адекватна побудована модель до експерименту, який розглядається, можна перевірити, як би це не звучало парадоксально, лише експериментально. Власне, здійснивши нескінченну кількість його повторень (що, звичайно ж, є неможливим). Якщо тепер припустити, що провели n — підкидань голки, а через m_n позначити кількість перетинів голкою прямих при цьому, то для частоти появи події $\nu_n = m_n/n$ повинна виконуватись рівність

$$\mathbb{P}(A) = \lim_{n\to+\infty} \nu_n = \frac{2\ell}{\pi a}, \qquad \text{де} \;\; \nu_n = \frac{m_n}{n}.$$

Звідси, для всіх достатньо великих n з наближеної рівності $\frac{2l}{\pi a} \approx \nu_n$ отримуємо наближену рівність

$$\pi \approx \frac{2\ell n}{a m_n},$$

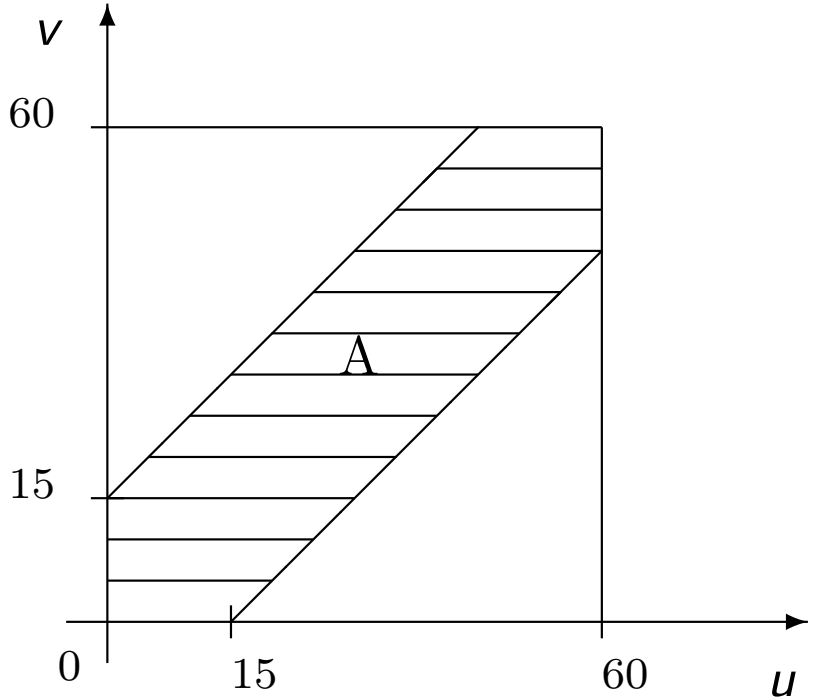
яку можна використати для перевірки адекватності моделі. Зауважимо, що якщо повірити в те, що побудована модель адекватно відображає експеримент, то на останню формулу можна подивитись і з точки зору джерела для відшукання числа π. Власне, таких експериментів для знаходження числа π за цією формулою різними дослідниками було проведено багато (див., наприклад, перелік на с. 36 у кн. [1]).

Вправа.

1. На круглий стіл радіуса a навмання кидаємо голку довжиною 2ℓ, $\ell < a$. Знайти ймовірність того, що: a) голка перетне край столу; b) голка впаде зі столу (тобто, центр ваги голки — її середина, — потрапить за край столу); c) розглянути попередні дві вправи, а також задачу Бюффона за умови $\ell \geq a$.

2. (задача про зустріч). Микола і Петро домовились зустрітись на зупинці автобуса між 9 і 10 годинами. Той, хто приходить першим, чекає 15 хвилин на іншого, а потім йде геть. Знайти ймовірність того, що зустріч відбудеться, вважаючи, що моменти приходу Миколи і Петра рівномірно розподілені в квадраті $\Omega := [9, 10] \times [9, 10]$.

Нехай $A = \{(u, v) \in \Omega \colon |u - v| \leq 0,25\}$ – подія, яка полягає в тому, що зустріч відбудеться. Тоді, $\mathbb{P}(A) = \frac{S(A)}{S(\Omega)} = \frac{S(\Omega) - 2S(\Delta)}{S(\Omega)}$, Δ – рівнобедрений прямокутний трикутник з катетами, що дорівнюють $0,75$. Звідки, $\mathbb{P}(A) = \frac{1 - 0,75^2}{1} = 0,4375$.



Нагадаємо означення зліченно-адитивної ймовірнісної міри.

Означення

Нехай $\mathcal{A} \subset 2^{\Omega}$ — деяка алгебра підмножин Ω.
Скінченно-адитивна ймовірнісна міра $\mathbb{P}(A)$, задана на $\mathcal{A}$, називається ймовірністю (зліченно-адитивною ймовірнісною мірою), якщо для будь-якої послідовності $(A_n)_{n=1}^{+\infty}$ такої, що $A_n \in \mathcal{A}\,(n \geqslant 1)$ і $A_n \cap A_k = \varnothing\,(n \neq k)$,

$$\bigcup_{n=1}^{+\infty} A_n \in \mathcal{A} \Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n).$$

Тепер нагадаємо означення ймовірності.

Означення

Зліченно-адитивну ймовірнісну міру, задану на σ-алгебрі називаємо ймовірністю.

1.5. Властивості ймовірності

1. Якщо $A \in \mathcal{A} \Rightarrow \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.

Доведення.

$\overline{A} \cup A = \Omega$, $A \cap \overline{A} = \varnothing$, $1 = \mathbb{P}(\Omega) = \mathbb{P}(\overline{A}) + \mathbb{P}(A)$. □

2. Нехай $A, B \in \mathcal{A}$ такі, що $B \subset A$. Тоді $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$.

Доведення.

$A \setminus B = A \cap \overline{B} = \overline{(\overline{A} \cup B)}$ (за формулою де Морґана). $\overline{A} \cap B = \varnothing$.

$$\mathbb{P}(A \setminus B) = 1 - \mathbb{P}(\overline{A} \cup B) = 1 - (\mathbb{P}(\overline{A}) + \mathbb{P}(B)) = \mathbb{P}(A) - \mathbb{P}(B).$$

□

3. (монотонність ймовірності). $B \subset A \Rightarrow \mathbb{P}(B) \leq \mathbb{P}(A)$.

Доведення.

$0 \leq \mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$. □

4. (напівадитивність ймовірності (узагальнена монотонність)).
$(\forall (A_j),\ A_j \in \mathcal{A}\ (j \in \mathbb{N}))$:

$$\mathbb{P}\Big(\bigcup_{j=1}^{+\infty} A_j\Big) \le \sum_{j=1}^{\infty} \mathbb{P}(A_j).$$

Доведення.

Позначимо $B_1 = A_1, B_2 = A_2 \cap \overline{A}_1, \ldots,\ B_n = A_n \cap \overline{A}_{n-1} \cap \overline{A}_{n-2} \cap \ldots \cap \overline{A}_1$.
Очевидно, що

$$B_j \cap B_k = \varnothing\ (j \ne k), \quad \bigcup_{j=1}^{+\infty} B_j = \bigcup_{j=1}^{+\infty} A_j, \quad B_j \subset A_j \quad (j \in \mathbb{N}).$$

Тому, за адитивністю і монотонністю ймовірності отримуємо

$$\mathbb{P}(\bigcup_{j=1}^{+\infty} A_j) = \mathbb{P}(\bigcup_{j=1}^{+\infty} B_j) = \sum_{j=1}^{+\infty} \mathbb{P}(B_j) \le \sum_{j=1}^{+\infty} \mathbb{P}(A_j).$$

$\square$

5. (теорема додавання). $(\forall N \in \mathbb{N})(\forall\ (A_j)_{j=1}^N,\ A_j \in \mathcal{A}\ (1 \le j \le N))$:

$$\mathbb{P}\Big(\bigcup_{j=1}^N A_j\Big) = \sum_{j=1}^N \mathbb{P}(A_j) - \sum_{1 \le j_1 < j_2 \le N} \mathbb{P}(A_{j_1} \cap A_{j_2}) + \ldots +$$
$$+(-1)^{k-1} \sum_{1 \le j_1 < j_2 < \ldots < j_k \le N} \mathbb{P}(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_k}) + \ldots +$$
$$+(-1)^{N-1}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n).$$

Доведення. Проведемо доведення за індукцією. Оскільки

$$A \cup B = A \setminus (A \cap B) + B \setminus (A \cap B) + A \cap B,$$

то для $n = 2$ маємо

$$\mathbb{P}(A \setminus (A \cap B)) + \mathbb{P}(B \setminus (A \cap B)) + \mathbb{P}(A \cap B) =$$
$$= \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B) =$$
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Доведемо, що можна зробити крок індукції. Нехай формула справджується для $n = m \geq 2$. Доведемо її для $n = m + 1$

$$\mathbb{P}\Big(\bigcup_{i=1}^{m+1} A_i\Big) = \mathbb{P}\Big(\Big(\bigcup_{i=1}^{m} A_i\Big) \cup A_{m+1}\Big) = \mathbb{P}\Big(\bigcup_{i=1}^{m} A_i\Big) + \mathbb{P}(A_{m+1}) -$$

$$- \mathbb{P}\Big(\bigcup_{i=1}^{m}(A_i \cap A_{m+1})\Big) = \Big\{\sum_{i=1}^{m} \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq m} \mathbb{P}(A_i \cap A_j) +$$

$$+ \sum_{1 \leq i < j < k \leq m} \mathbb{P}(A_i \cap A_j \cap A_k) - \ldots + (-1)^{m-1}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_m)\Big\} +$$

$$+ \mathbb{P}(A_{m+1}) - \Big\{\sum_{i=1}^{m} \mathbb{P}(A_i \cap A_{m+1}) - \sum_{1 \leq i < j \leq m} \mathbb{P}(A_i \cap A_j \cap A_{m+1}) + \ldots$$

$$+ (-1)^{m-1}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{m+1})\Big\} = \sum_{i=1}^{m+1} \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq m+1} \mathbb{P}(A_i \cap A_j) +$$

$$+ \sum_{1 \leq i < j < k \leq m+1} \mathbb{P}(A_i \cap A_j \cap A_k) - \ldots + (-1)^{m}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{m+1}).$$

Теорема (теорема неперервності ймовірності)

Нехай $\mathbb{P}$ — скінченно-адитивна ймовірнісна міра на $\mathcal{A}$, $\mathcal{A}$ — σ-алгебра. Наступні три твердження еквівалентні:

1. $\mathbb{P}$ — σ-адитивна міра, тобто, ймовірність.
2. $\big(\forall (A_n),\ A_n \in \mathcal{A},\ A_n \subset A_{n+1}\ (n \geq 1)\big):\ \mathbb{P}\big(\bigcup\limits_{n=1}^{+\infty} A_n\big) = \lim\limits_{n\to+\infty} \mathbb{P}(A_n).$
3. $\big(\forall (B_n),\ B_n \in \mathcal{A}, B_n \supset B_{n+1}\ (n \geq 1)\big):\ \mathbb{P}\big(\bigcap\limits_{n=1}^{+\infty} B_n\big) = \lim\limits_{n\to+\infty} \mathbb{P}(B_n).$

Доведення. Еквівалентність 2. $\Leftrightarrow$ 3. випливає з того, що

$$A_n \subset A_{n+1}\ (n \geq 1)\ \Leftrightarrow\ B_{n+1} \subset B_n\ (n \geq 1), \quad \text{де } B_n = \overline{A}_n\ (n \geq 1),$$

$$\bigcap_{n=1}^{+\infty} B_n = \bigcap_{n=1}^{+\infty} \overline{A}_n = \overline{\big(\bigcup_{n=1}^{+\infty} A_n\big)}, \quad \mathbb{P}\big(\bigcap_{n=1}^{+\infty} B_n\big) = 1 - \mathbb{P}\big(\bigcup_{n=1}^{+\infty} A_n\big),$$

$$\lim_{n\to+\infty} \mathbb{P}(B_n) = 1 - \lim_{n\to+\infty} \mathbb{P}(A_n).$$

1. $\Rightarrow$ 2. Нехай $A_n \subset A_{n+1}$ $(n \geq 1)$. Розглянемо послідовність попарно неперетинних множин

$$C_1 = A_1, \quad C_n = A_n \setminus A_{n-1} \ (n \geq 2).$$

Зрозуміло, що $\bigcup\limits_{n=1}^{+\infty} C_n = \bigcup\limits_{n=1}^{+\infty} A_n$. Тому, з огляду на зліченну адитивність $\mathbb{P}$

$$\mathbb{P}\Big(\bigcup_{n=1}^{+\infty} A_n\Big) = \mathbb{P}\Big(\bigcup_{n=1}^{+\infty} C_n\Big) = \sum_{k=1}^{+\infty} \mathbb{P}(C_k) =$$

$$= \mathbb{P}(A_1) + \sum_{k=2}^{+\infty} (\mathbb{P}(A_k) - \mathbb{P}(A_{k-1})) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

2. $\Rightarrow$ 1. Нехай (C_n) послідовність попарно неперетинних множин. Розглянемо монотонно зростаючу послідовність множин $A_1 = C_1$, $A_n = \bigcup\limits_{k=1}^{n} C_n$ $(n \geq 2)$. Оскільки $\bigcup\limits_{n=1}^{+\infty} C_n = \bigcup\limits_{n=1}^{+\infty} A_n$, то зі скінченної адитивності $\mathbb{P}$ маємо

$$\mathbb{P}\Big(\bigcup_{n=1}^{+\infty} C_n\Big) = \mathbb{P}\Big(\bigcup_{n=1}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mathbb{P}(A_n) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcup_{k=1}^{n} C_k\Big) \Longleftrightarrow$$

$$= \lim_{n\to+\infty} \sum_{k=1}^{n} \mathbb{P}(C_k) = \sum_{k=1}^{+\infty} \mathbb{P}(C_k).$$

1.7. Умовні ймовірності

Розглянемо спочатку такий

Приклад. Підкидаємо двічі кубик. $\Omega = \{(i,j)\colon 1 \le i,j \le 6\}$. Подія $A = \{(i,j)\colon i+j \le 4\}$ — сума очок, яка випала, не більша за 4. Подія B — при першому підкиданні випала одиничка $B = \{(1,j)\colon 1 \le j \le 6\}$.
Ймовірності цих подій $\mathbb{P}(A) = \frac{6}{36} = \frac{1}{6}$, $\mathbb{P}(B) = \frac{6}{36} = \frac{1}{6}$.
Через $\mathbb{P}(A|B)$ позначимо ймовірність події A за умови, що подія B відбулася. У даному експерименті — це ймовірність того, що сума очок, які випали при двох підкиданнях, не перевищить 4, за умови, що при першому підкиданні випала одиниця. Очевидно, що
$\mathbb{P}(A|B) = \mathbb{P}(C) = \frac{3}{6} = \frac{1}{2}$, де $C = \{(1,j) : 1 \le j \le 3\}$.
Зауважимо, що ймовірність добутку подій $\mathbb{P}(A \cap B) = \frac{3}{36} = \frac{1}{12}$, де $A \cap B = \{(i,j) : i = 1 \wedge i+j \le 4\}$. Тому, $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

Твердження

Нехай в рамках класичної схеми B — подія з додатною ймовірністю $\mathbb{P}(B) > 0$. Тоді, ймовірність події A за умови, що подія B відбулася дорівнює

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

Вправа. Довести це твердження.
Інтуїтивно зрозуміло, що для геометричних ймовірностей ця формула також повинна бути правильною. В загальному, умовну ймовірність події A за умови, що подія B з ймовірністю $\mathbb{P}(B) > 0$ відбулася, визначимо за цією формулою. Пізніше, розглядаючи умовні розподіли, ми повернемось до поняття умовної ймовірності у загальній ситуації.

Означення (умовної ймовірності)

Нехай $(\Omega, \mathcal{A}, \mathbb{P})$ — ймовірнісний простір, а $B \in \mathcal{A}$ — подія з додатною ймовірністю $\mathbb{P}(B) > 0$. Умовною ймовірністю події A за умови, що подія B відбулася, називається число

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

Введемо новий (умовний) ймовірнісний простір. Нехай $B \in \mathcal{A}$ — фіксована подія з додатною ймовірністю $\mathbb{P}(B) > 0$. Через $\mathcal{A}_B = \{A \cap B : A \in \mathcal{A}\}$ позначимо σ-алгебру, яку отримаємо з $\mathcal{A}$, перетнувши кожен її елемент з B. Умовна ймовірність $\mathbb{P}_B(A) \stackrel{def}{=} \mathbb{P}(A|B)$.
Вправа. Перевірити, що трійка $(B, \mathcal{A}_B, \mathbb{P}_B)$ — ймовірнісний простір (це й буде умовний ймовірнісний простір).

1.9. Формула множення ймовірностей

Твердження

Нехай $(A_j)_{j=1}^n$, $A_j \in \mathcal{A}$, $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$. Тоді

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \times \ldots$$
$$\ldots \times \mathbb{P}(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \text{ — формула множення.}$$

Доведення. При $n = 2$ з означення умовної ймовірності маємо

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2)\mathbb{P}(A_1|A_2).$$

Припущення індукції. Формула правильна для числа подій не більшого за $n-1$. Доведемо, що вона правильна для кількості подій, що дорівнює n. Тоді за методом математичної індукції отримаємо, що формула правильна для всіх $n \in \mathbb{N}$.

Крок індукції. Виберемо $A = A_n$, $B = \bigcap_{j=1}^{n-1} A_j$ і застосуємо для цих двох подій формулу множення

$$\mathbb{P}(\bigcap_{j=1}^{n} A_j) = \mathbb{P}(A \cap B) = \mathbb{P}(\bigcap_{j=1}^{n-1} A_j) \cdot \mathbb{P}(A_n|\bigcap_{j=1}^{n-1} A_j).$$

Але за припущенням індукції

$$\mathbb{P}(\bigcap_{j=1}^{n-1} A_j) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\ldots\mathbb{P}(A_{n-1}|A_1 \cap A_2 \cap \ldots A_{n-2}).$$

Ще слід пояснити, чому всі записані умовні ймовірності існують. Справді, за умовою і монотонністю ймовірності, оскільки

$$A_1 \supset A_1 \cap A_2 \supset (A_1 \cap A_2 \cap A_3) \supset \ldots \supset (A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

то $\mathbb{P}(\bigcap_{j=1}^{n-1} A_j) > 0 \Rightarrow \mathbb{P}(\bigcap_{j=1}^{n-2} A_j) > 0, \ldots, \mathbb{P}(A_1 \cap A_2) > 0, \mathbb{P}(A_1) > 0.$ Отже, все визначено коректно.

Приклад

З ящика, що містить 3 білих і 2 чорні кулі, за схемою випадкового вибору без повернення послідовно виймають кулі. Знайти ймовірність p_k того, що чорна куля вперше з'явиться при k-му випробуванні.

Нехай C_i, $i \in \{1, 2, 3, 4\}$, подія, яка полягає в тому, що в i-му випробуванні з'явилась чорна куля, $B_k = \{$чорна куля вперше з'явилась при k-му випробуванні$\}$, $k \in \{1, 2, 3, 4\}$. Тоді $B_1 = C_1$, $B_2 = \overline{C}_1 \cap C_2$, $B_3 = \overline{C}_1 \cap \overline{C}_2 \cap C_3$, $B_4 = \overline{C}_1 \cap \overline{C}_2 \cap \overline{C}_3 \cap C_4$. За формулою множення ймовірностей

$$\mathbb{P}(B_1) = \mathbb{P}(C_1), \quad \mathbb{P}(B_2) = \mathbb{P}(\overline{C}_1)\mathbb{P}(C_2|\overline{C}_1),$$

$$\mathbb{P}(B_3) = \mathbb{P}(\overline{C}_1)\mathbb{P}(\overline{C}_2|\overline{C}_1)\mathbb{P}(C_3|\overline{C}_1 \cap \overline{C}_2),$$

$$\mathbb{P}(B_4) = \mathbb{P}(\overline{C}_1)\mathbb{P}(\overline{C}_2|\overline{C}_1)\mathbb{P}(\overline{C}_3|\overline{C}_1 \cap \overline{C}_2)\mathbb{P}(C_4|\overline{C}_1 \cap \overline{C}_2 \cap \overline{C}_3).$$

За класичним означенням ймовірності

$$\mathbb{P}(C_1) = \frac{2}{5}, \ \mathbb{P}(\overline{C}_1) = \frac{3}{5}, \ \mathbb{P}(\overline{C}_{i+1}|\overline{C}_1 \cap \dots \cap \overline{C}_i) = \frac{3-i}{5-i},$$

$$\mathbb{P}(C_{i+1}|\overline{C}_1 \cap \dots \cap \overline{C}_i) = \frac{2}{5-i}.$$

У результаті $p_1 = \mathbb{P}(B_1) = 0,4$, $p_2 = \mathbb{P}(B_2) = 0,3$, $p_3 = \mathbb{P}(B_3) = 0,2$, $p_4 = \mathbb{P}(B_4) = 0,1$.

1.10. Повна група гіпотез та формула повної ймовірності

Розглянемо ймовірнісний простір $(\Omega, \mathcal{A}, \mathbb{P})$. Нехай $(H_j)_{j=1}^N$ $(N \le +\infty)$ — скінченний або зліченний набір подій. Система (H_j) називається повною групою подій (повною групою гіпотез), якщо виконуються такі умови:

1) $\{H_j\}_{j=1}^N$ — попарно несумісні події, тобто $(\forall j \ne k)$: $H_j \cap H_k = \varnothing$;

2) $\bigcup\limits_{j=1}^{N} H_j = \Omega$; 3) $\mathbb{P}(H_j) > 0$ $(j \ge 1)$.

Для довільної події $A \in \mathcal{A}$, скориставшись означенням повної групи подій, можемо записати

$$A = A \cap \Omega = A \cap (\bigcup_{j=1}^{N} H_j) = \bigcup_{j=1}^{N} (A \cap H_j).$$

Звідси, оскільки, $(\forall j \ne k)$: $(A \cup H_j) \cap (A \cup H_k) = \varnothing$, за зліченною адитивністю ймовірності, скориставшись формулою для ймовірності добутку подій, отримуємо формулу повної ймовірності

$$\mathbb{P}(A) = \sum\nolimits_{j=1}^{N} \mathbb{P}(A \cap H_j) = \sum\nolimits_{j=1}^{N} \mathbb{P}(H_j)\mathbb{P}(A|H_j).$$

Всі записані ймовірності гіпотез $\mathbb{P}(H_j)$ — апріорні ймовірності, тобто, ймовірності до початку експерименту.

Приклад

У скриньку, що містить 8 стандартних виробів, додають 2 вироби зі складу; відомо, що частка бракованих виробів на складі дорівнює 5%. Знайти ймовірність того, що навмання взятий з поповненої скриньки виріб не буде бракованим.

Означимо події A, H_0, H_1, H_2: $A = \{$виріб, взятий з поповненої скриньки, стандартний$\}$, $H_k = \{$з двох виробів, доданих у скриньку, k бракованих$\}$, $k \in \{0, 1, 2\}$.
Зрозуміло, що $H_0 \cup H_1 \cup H_2 = \Omega$ і $H_k \cap H_l = \varnothing$ $(k \neq l)$, тобто ці події утворюють повну групу подій. Скористаємось формулою повної ймовірності $\mathbb{P}(A) = \mathbb{P}(H_0)\mathbb{P}(A|H_0) + \mathbb{P}(H_1)\mathbb{P}(A|H_1) + \mathbb{P}(H_2)\mathbb{P}(A|H_2)$.
Якщо відбулась подія H_k, то в ящику з 10 виробів k бракованих, отже, $\mathbb{P}(A \mid H_k) = \frac{10-k}{10}$, $k \in \{0, 1, 2\}$.
При виборі малої кількості виробів з великої партії схеми вибору з поверненням і без повернення мають близькі ймовірності результатів. Тому вважаємо, що кожен з доданих виробів незалежно один від одного може бути бракованим з ймовірністю 0,05. Отже, $\mathbb{P}(H_0) = 0,95^2$, $\mathbb{P}(H_1) = 2 \cdot 0,95 \cdot 0,05$, $\mathbb{P}(H_2) = 0,05^2$. Остаточно,

$$\mathbb{P}(A) = 0,95^2 \cdot 1 + 2 \cdot 0,95 \cdot 0,05 \cdot 0,9 + 0,05^2 \cdot 0,8 = 0,99.$$

1.11. Приклад (кульки і скриньки) та формула Баєса

Маємо N скриньок з однаковими за розміром кульками. В j-тій скринці n_j кульок, з яких m_j білих кульок. Експеримент полягає у випадковому виборі з навмання вибраної скриньки кульки. Подія A — вийнята кулька виявилася білою. Знайдемо ймовірність цієї події.
Через H_j позначимо подію — навмання вибрано j-ту скриньку. Якщо припустити, що вибір скриньок — однаково можливі події, то $\mathbb{P}(H_j) = \frac{1}{N}$; $\mathbb{P}(A|H_j) = \frac{m_j}{n_j}$ — умовна ймовірність події "вийняти білу кульку з j-тої скриньки". Тоді, за формулою повної ймовірності

$$\mathbb{P}(A) = \frac{1}{N} \sum\nolimits_{j=1}^{N} \frac{m_j}{n_j}.$$

Припустимо тепер, що після проведення експерименту виявилось, що вийнята куля — білого кольору. Тоді, спочатку застосовуючи формулу для умовної ймовірності, а потім формулу повної ймовірності, отримаємо

$$\mathbb{P}(H_j|A) = \frac{\mathbb{P}(H_j \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(H_j)\mathbb{P}(A|H_j)}{\sum_{j=1}^{k} \mathbb{P}(H_j)\mathbb{P}(A|H_j)}.$$

$$\mathbb{P}(H_j|A) = \frac{\mathbb{P}(H_j)\mathbb{P}(A|H_j)}{\sum_{j=1}^{k} \mathbb{P}(H_j)\mathbb{P}(A|H_j)} \quad \text{— формула Баєса.}$$

$\mathbb{P}(H_j|A)$ — це апостеріорні ймовірності гіпотез, тобто ймовірності знайдені після проведення експерименту.

Остання формула є правильною для довільної повної групи подій і довільної події A з $\mathbb{P}(A) > 0$. Для нашого експерименту з кульками ймовірність того, що білу кульку вийнято з j-тої скриньки дорівнює

$$\mathbb{P}(H_j|A) = \frac{\frac{1}{N}\frac{m_j}{n_j}}{\frac{1}{N}\sum_{k=1}^{N}\frac{m_k}{n_k}} = \frac{m_j}{n_j\sum_{k=1}^{N}\frac{m_k}{n_k}}.$$

Зрозуміло, як зміняться знайдені формули у випадку, коли вибір скриньок не є однаково можливим, тобто, $\mathbb{P}(H_j) = p_j$, $\sum_{j=1}^{N} p_j = 1$. Тоді

$$\mathbb{P}(A) = \sum_{j=1}^{N}\frac{m_j}{n_j}p_j, \quad \mathbb{P}(H_j|A) = \frac{m_j p_j}{n_j\sum_{k=1}^{N}\frac{m_k}{n_k}p_k}.$$

Звідси, зокрема, випливає, що з найбільшою ймовірністю білу кульку вийнято зі скриньки для якої відносна “вага” білих кульок $\frac{m_j}{n_j}$ є найбільшою.

Приклад (щасливі білети)

На іспит викладач підготував n екзаменаційних білетів, серед яких є $m \leq n$ — "щасливих". Двоє студентів один за одним навмання вибирають білети. У кого з них більше шансів вийняти “щасливий” білетї

Розглянемо події A_j — j-тий студент вийняв щасливий білет, $j \in \{1, 2\}$. Зрозуміло, що $\mathbb{P}(A_1) = \frac{m}{n}$. Для знаходження такої ж ймовірності для студента, який зайшов другим розглянемо таку повну групу гіпотез: $H_1 = A_1$ — перший студент витягнув щасливий білет, $H_2 = \overline{H}_1$ — перший студент не витягнув щасливий білет. Тоді, за формулою повної ймовірності для $1 \leq m \leq n$ отримаємо

$$\mathbb{P}(A_2) = \mathbb{P}(H_1) \cdot \mathbb{P}(A_2|H_1) + \mathbb{P}(H_2)\mathbb{P}(A_2|H_2) =$$
$$= \frac{m}{n}\frac{m-1}{n-1} + \frac{n-m}{n}\frac{m}{n-1} = \frac{m(n-1)}{n(n-1)} = \frac{m}{n}.$$

Тобто, перед початком іспиту в обидвох студентів шанси вибрати щасливий білет однакові.

Вправа. Обчислити ймовірність вийняти “щасливий” білет для: а) третього студента; б) k-того ($k \leq m$) студента.

Приклад

Зі скриньки, в якій було 7 білих і 5 червоних куль, випадково загубили 2 кулі. Після цього зі скриньки навмання вийняли одну кулю, яка виявилась червоною. Знайти ймовірність того, що загублено 2 білі кулі.

Розглянемо гіпотези H_i — серед загублених куль i білих, $i \in \{0, 1, 2\}$ та подію A — після втрат куль зі скриньки вийняли червону кулю.
$\{H_0, H_1, H_2\}$ — повна група подій, тому можемо за формулою Баєса знайти шукану апостеріорну ймовірність

$$\mathbb{P}(H_2|A) = \frac{\mathbb{P}(H_2)\mathbb{P}(A|H_2)}{\mathbb{P}(H_0)\mathbb{P}(A|H_0) + \mathbb{P}(H_1)\mathbb{P}(A|H_1) + \mathbb{P}(H_2)\mathbb{P}(A|H_2)} = \frac{21}{55},$$

де за класичним означенням ймовірності

$$\mathbb{P}(H_0) = \frac{C_5^2}{C_{12}^2} = \frac{20}{132}, \ \mathbb{P}(H_1) = \frac{C_5^1 C_7^1}{C_{12}^2} = \frac{70}{132}, \ \mathbb{P}(H_2) = \frac{C_7^2}{C_{12}^2} = \frac{42}{132},$$

$\mathbb{P}(A|H_0) = \frac{3}{10}$ — ймовірність того, що у скриньці залишилось 7 білих і 3 червоні кулі,
$\mathbb{P}(A|H_1) = \frac{4}{10}$ — ймовірність того, що у скриньці залишилось 6 білих і 4 червоні кулі,
$\mathbb{P}(A|H_2) = \frac{5}{10}$ — ймовірність того, що у скриньці залишилось 5 білих і 5 червоних куль.

Приклад

У скриньці розташована куля, про яку відомо, що вона чорна або біла. В урну поклали білу кулю, а потім після ретельного перемішування вийняли кулю, яка виявилась білою. Яка ймовірність того, що після цього виймуть з урни білу кулю? (Льюис Кэррол. История с узелками. М., Мир, 1973)

На перший погляд здається, що нічого і не відбулось, відповідь — $1/2$. Але насправді перший раз біла куля була вийнята випадково, тому задача потребує глибшого ймовірнісного аналізу, який приводить до іншої відповіді.

Експеримент складається з двох частин — перше виймання кулі і друге виймання кулі. Події: A — перша вийнята куля виявилась білою, B — друга вийнята куля виявилась білою. Нехай $\mathbb{P}(H_i)$ — апріорні ймовірності гіпотез H_1 — в урні була біла куля, H_2 — в урні була чорна куля. За результатом першої частини експерименту ми можемо переоцінити ці ймовірності — отримаємо апостеріорні ймовірності $\mathbb{P}(H_i|A)$, які відіграють роль апріорних для другої частини експерименту.

Застосовуючи формули Баєса до першої частини і формулу повної ймовірності до другої, маємо

$$\mathbb{P}(H_i) = \frac{1}{2}, \quad i \in \{1, 2\}, \quad \mathbb{P}(H_1|A) = \frac{\frac{1}{2} \cdot 1}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2}} = \frac{2}{3},$$

$$\mathbb{P}(H_2|A) = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2}} = \frac{1}{3},$$

$$\mathbb{P}(B) = \sum_{i=1}^{2} \mathbb{P}(H_i|A)\mathbb{P}(B|(H_i|A)) = \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot 0 = \frac{2}{3}.$$

1.12. Незалежні події

Означення (незалежних подій)

Події $A, B \in \mathcal{A}$ називають незалежними, якщо
$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Твердження

Нехай $A, B \in \mathcal{A}$, $\mathbb{P}(B) > 0$. Події A і B — незалежні $\Longleftrightarrow \mathbb{P}(A|B) = \mathbb{P}(A)$.

Вправа. Довести це твердження самостійно.

Кажуть, що $(A_j)_{j=1}^N$ $(N \leq +\infty)$ — послідовність попарно незалежних подій, якщо A_j і A_k — незалежні $(\forall j \neq k)$.

Означення (незалежності в сукупності)

Події $(A_j)_{j=1}^N, N \leq +\infty$ — незалежні в сукупності, якщо

$$(\forall n \in \mathbb{N})(\forall 1 \leq i_1 < i_2 \leq \ldots i_n \leq N)\colon \ \mathbb{P}\Big(\bigcap_{k=1}^{n} A_{i_k}\Big) = \prod_{k=1}^{n} \mathbb{P}(A_{i_k}).$$

Надалі, під незалежністю скрізь розуміємо незалежність в сукупності.

Приклад (С. Н. Бернштейна). Підкидають правильну трикутну піраміду (тетраедр), кожна з граней якої розфарбована відповідно у один із кольорів: синій, жовтий, зелений, а основа — у всі три кольори. Скажемо, що при підкиданні піраміди випав певний колір, якщо вона впала на грань, у розфарбуванні якої використано цей колір.
Подію, що означає випадання певного кольору позначаємо першою літерою з назви кольору: ж – жовтий, з – зелений, с – синій. Тоді,

$$\mathbb{P}(\text{ж}) = \mathbb{P}(\text{с}) = \mathbb{P}(\text{з}) = \frac{2}{4} = \frac{1}{2}, \quad \mathbb{P}(\text{ж} \cap \text{с}) = \frac{1}{4} = \mathbb{P}(\text{ж})\mathbb{P}(\text{с}),$$

$$\mathbb{P}(\text{ж} \cap \text{з}) = \frac{1}{4} = \mathbb{P}(\text{ж})\mathbb{P}(\text{з}), \quad \mathbb{P}(\text{с} \cap \text{з}) = \frac{1}{4} = \mathbb{P}(\text{с})\mathbb{P}(\text{з}) =,$$

$$\mathbb{P}(\text{с} \cap \text{з} \cap \text{ж}) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(\text{ж})\mathbb{P}(\text{с})\mathbb{P}(\text{з}),$$

тобто, події ж, с, з незалежні попарно, але не є незалежні в сукупності.
Очевидно, що правильне таке твердження.

Твердження

Якщо події (A_n) незалежні в сукупності, то вони — незалежні попарно.

Як показує приклад С.Н. Бернштейна, зворотна імплікація —

Вправа. Для кожного $n \in \mathbb{N}$, $n \geq 5$, побудувати приклад послідовності подій $A_1, A_2, \ldots, A_n$ таких, що

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_{n-1}}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_{n-1}})$$

для кожного набору індексів $1 \leq i_1 < i_2 < \cdots < i_{n-1} \leq n$, але

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) \neq \mathbb{P}(A_1)\mathbb{P}(A_2) \cdots \mathbb{P}(A_n).$$

1.13. Незалежність σ-алгебр та експериментів

Нехай є набір ймовірнісних просторів: $(\Omega, \mathcal{A}_j, \mathbb{P})$, де $\mathcal{A}_j \subset \mathcal{A}$ — σ-алгебри.

Означення (незалежності σ-алгебр)

Скажемо, що $(\mathcal{A}_j)$ — послідовність незалежних σ-алгебр, якщо $(\forall (A_j),\ A_j \in \mathcal{A}_j)$: (A_j) — послідовність незалежних в сукупності подій.

Оскільки з кожним експериментом ми пов'язуємо, взагалі кажучи, свій ймовірнісний простір, зокрема, свою σ-алгебру, то поняття незалежності σ-алгебр природно породжує поняття незалежності послідовності експериментів.

Означення

Скажемо, що послідовність експериментів є послідовністю незалежних експериментів, якщо незалежні їхні σ-алгебри.

Приклад

Ймовірність того, що в результаті чотирьох незалежних випробувань подія A настане принаймні один раз, дорівнює 0,59. Знайти ймовірність настання події A при одному випробуванні, якщо вона однакова під час всіх випробувань.

Нехай $A_i = \{\text{подія } A \text{ настала при } i\text{-му випробуванні}\}$, $i \in \{1, 2, 3, 4\}$, тоді треба знайти $p = \mathbb{P}(A_i)$. За умовою $\mathbb{P}(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \overline{A}_4) = 1 - 0,59 = 0,41$. Оскільки події A_i незалежні, то $\overline{A}_i$ також незалежні, $i \in \{1, 2, 3, 4\}$. Тому,

$$\mathbb{P}(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \overline{A}_4) = \mathbb{P}(\overline{A}_1)\mathbb{P}(\overline{A}_2)\mathbb{P}(\overline{A}_3)\mathbb{P}(\overline{A}_4) = (1-p)^4,$$

звідки $p = 1 - \sqrt[4]{0,41}$. Залишається зауважити, що наприклад для двох подій маємо $\mathbb{P}(\overline{A}_1 \cap \overline{A}_2) = \mathbb{P}(\overline{A_1 \cup A_2}) = 1 - \mathbb{P}(A_1 \cup A_2) =$
$= 1 - \big(\mathbb{P}(A_1) + \mathbb{P}(A_2) - P(A_1)\mathbb{P}(A_2)\big) = (1 - \mathbb{P}(A_1))(1 - \mathbb{P}(A_2)) = \mathbb{P}(\overline{A}_1)\mathbb{P}(\overline{A}_2)$.
І, подібно, в загальному

$$\mathbb{P}(\overline{A}_1 \cap \ldots \cap \overline{A}_n) = \mathbb{P}(\overline{A_1 \cup \ldots \cup A_n}) = 1 - \mathbb{P}(A_1 \cup \ldots \cup A_n) =$$

$$= 1 - \Big(\sum \mathbb{P}(A_k) - \sum_{k<l} \mathbb{P}(A_k)\mathbb{P}(A_l) + \sum_{k<l<m} \mathbb{P}(A_k)\mathbb{P}(A_l)\mathbb{P}(A_m) - \ldots\Big) =$$

$$= \prod \big(1 - \mathbb{P}(A_k)\big) = \prod \mathbb{P}(\overline{A}_k),$$

тобто, $(\overline{A}_k)$ – незалежні події як тільки (A_k) – незалежні події.

2.1. Випадкові величини і випадкові вектори, незалежність випадкових величин

Пригадаємо, що функція $f : \mathbb{R}^m \to \mathbb{R}^n$ $(m, n \in \mathbb{N})$ називається борелевою, якщо прообраз будь-якої борелевої множини на $\mathbb{R}^n$ є борелевою множиною в $\mathbb{R}^m$, тобто,

$$(\forall B \in \mathcal{B}(\mathbb{R}^n)) : f^{-1}(B) \stackrel{def}{=} \{\omega : f(\omega) \in B\} \in \mathcal{B}(\mathbb{R}^m).$$

Нехай $(\Omega, \mathcal{A}, \mathbb{P})$ — ймовірнісний простір. Функцію $g : \Omega \to \mathbb{R}$ називатимемо вимірною (відносно пари σ-алгебр $(\mathcal{B}(\mathbb{R}), \mathcal{A})$), якщо для кожної борелевої множини її прообраз є подією (елементом з $\mathcal{A}$), тобто
$(\forall B \in \mathcal{B}(\mathbb{R}^n)) : g^{-1}(B) \in \mathcal{A}$

Означення (випадкової величини)

Кожну вимірну функцію відносно пари σ-алгебр $(\mathcal{B}(\mathbb{R}), \mathcal{A})$ називають випадковою величиною.

Власне, ξ – випадкова величина, якщо для будь-якої борелевої множини $B \in \mathcal{B}(\mathbb{R})$ її прообраз

$$\xi^{-1}(B) \stackrel{def}{=} \{\omega \in \Omega \colon \xi(\omega) \in B\} \in \mathcal{A},$$

тобто, є подією.

Приклад

1. Нехай $A \subset \Omega$, $A \notin \mathcal{A}$. Тоді, характеристична функція (індикатор) множини A, $\chi_A(\omega) = \begin{cases} 1 & \omega \in A, \\ 0 & \omega \notin A, \end{cases}$ — не вимірна відносно пари σ-алгебр $(\mathcal{B}(\mathbb{R}), \mathcal{A})$, а, отже, не випадкова величина.

Справді, множина $B = \{1\}$ – борелева, а $\chi_A^{-1}(B) = A \notin \mathcal{A}$.

2. Нехай $\mathcal{A}_\infty = 2^\Omega$. Тоді кожна функція $f : \Omega \to \mathbb{R}$ — випадкова величина.

Скрізь надалі будемо використовувати такі загальні властивості операції "взяття прообразу".

Твердження (взяття прообразу зберігає теоретико-множинні операції)

Нехай $\xi\colon \mathcal{B}_1 \to \mathcal{B}_2$ — деяке відображення з простору $\mathcal{B}_1$ у простір $\mathcal{B}_2$. Тоді $(\forall (B_\alpha)_{\alpha\in I},\ B_\alpha \subset \mathcal{B}_2)$:

1) $\xi^{-1}(\bigcup\limits_{\alpha\in I} B_\alpha) = \bigcup\limits_{\alpha\in I} \xi^{-1}(B_\alpha)$,

2) $\xi^{-1}(\bigcap\limits_{\alpha\in I} B_\alpha) = \bigcap\limits_{\alpha\in I} \xi^{-1}(B_\alpha)$,

3) $\xi^{-1}(B_{\alpha_1} \setminus B_{\alpha_2}) = \xi^{-1}(B_{\alpha_1}) \setminus \xi^{-1}(B_{\alpha_2})$ $(\alpha_j \in I)$,

4) $\xi^{-1}(\mathcal{B}_2) = \mathcal{B}_1$,

5) $\xi^{-1}(\varnothing) = \varnothing$.

Доведемо, наприклад, п. 1). Маємо, $\omega \in \xi^{-1}(\bigcup\limits_{\alpha\in I} B_\alpha) \Longleftrightarrow$ $\xi(\omega) \in \bigcup\limits_{\alpha\in I} B_\alpha \Longleftrightarrow (\exists \alpha \in I)\colon\ \xi(\omega) \in B_\alpha \Longleftrightarrow$ $(\exists \alpha \in I)\colon\ \omega \in \xi^{-1}\Big(B_\alpha\Big) \Longleftrightarrow \omega \in \bigcup\limits_{\alpha\in I} \xi^{-1}(B_\alpha)$,

Вправа. Довести п.п. 2)–5) цього твердження.

Розглянемо $(\Omega, \mathcal{A}, \mathbb{P})$ — ймовірнісний простір і функцію $\xi = \xi(\omega)\colon \Omega \to \mathbb{R}$. Наступне твердження дає простий критерій випадкової величини.

Твердження (критерій випадкової величини)

Функція $\xi\colon \Omega \to \mathbb{R}$ є випадковою величиною тоді і тільки тоді, коли

$$(\forall x \in \mathbb{R})\colon\ \xi^{-1}((-\infty, x)) \stackrel{def}{=} \{\omega \in \Omega\colon \xi(\omega) < x\} \in \mathcal{A}.$$

Доведення. Необхідність. Оскільки, нескінченний інтервал $(-\infty, x) = \bigcup\limits_{n=1}^{+\infty}[-n, x)$, $x \in \mathbb{R}$ є борелевою множиною, то, вибираючи $B = (-\infty, x)$, необхідність умови отримуємо за означенням.

Достатність. Розглянемо сім'ю множин $\mathcal{M} = \{B \subset \mathbb{R}\colon \xi^{-1}(B) \in \mathcal{A}\}$.

Якщо ми доведемо, що $\mathcal{M}$ — σ-алгебра, яка містить всі півінтервали вигляду $[a, b)$, то з того, що σ-алгебра борелевих множин на прямій є мінімальною σ-алгеброю, породженою півінтервалами зазначеного вигляду, отримуємо, що

Отже, оскільки за умовою $\xi^{-1}((-\infty, x)) \in \mathcal{A}$ для кожного x, а $\mathcal{A}$ — σ-алгебра, то для $a < b$ послідовно маємо

$$\xi^{-1}([a, b)) = \xi^{-1}((-\infty, b) \setminus (-\infty, a)) =$$
$$= \xi^{-1}((-\infty, b)) \setminus \xi^{-1}((-\infty, a)) \in \mathcal{A},$$

Тому $[a, b) \in \mathcal{M}$.
Зауважимо тепер, що $\xi^{-1}(\mathbb{R}) = \Omega \in \mathcal{A}$, звідки $\mathbb{R} \in \mathcal{M}$.
Тоді, оскільки для кожної $B \in \mathcal{M}$ також

$$\xi^{-1}(\mathbb{R} \setminus B) = \xi^{-1}(\mathbb{R}) \setminus \xi^{-1}(B) = \Omega \setminus \xi^{-1}(B) \in \mathcal{A},$$

то $\mathbb{R} \setminus B \in \mathcal{M}$.
І, нарешті, для кожної послідовності (B_j), $B_j \in \mathcal{M}$ $(j \geq 1)$ маємо $\xi^{-1}(\bigcup\limits_{n=1}^{+\infty} B_n) = \bigcup\limits_{n=1}^{+\infty} \xi^{-1}(B_n) \in \mathcal{A}$, тобто, $\bigcup\limits_{n=1}^{+\infty} B_n \in \mathcal{M}$.

Означення (випадкового вектора)

Функцію $\xi = (\xi_1, \ldots, \xi_n)\colon \Omega \to \mathbb{R}^n$ називатимемо випадковим вектором, якщо всі ξ_j $(1 \le j \le n)$ — випадкові величини.

Надалі випадкові величини і випадкові вектори позначатимемо літерами грецького алфавіту $\xi, \eta, \zeta, \ldots$.
Через

$$\mathcal{A}_\xi = \xi^{-1}(\mathcal{B}(\mathbb{R})) = \{\xi^{-1}(B)\colon B \in \mathcal{B}(\mathbb{R})\}$$

позначимо σ-алгебру, породжену випадковою величиною ξ.

Вправа. Переконатись, що $\mathcal{A}_\xi$ справді — σ-алгебра.

Означення (незалежних випадкових величин)

Випадкові величини ξ_1, ξ_2 називаються незалежними, якщо σ-алгебри $\mathcal{A}_{\xi_1}, \mathcal{A}_{\xi_2}$, породжені цими випадковими величинами, є незалежними.

Означення (послідовності незалежних випадкових величин)

Послідовність випадкових величин (ξ_j) є послідовністю незалежних випадкових величин, якщо $(\mathcal{A}_{\xi_j})$ — послідовність незалежних σ-алгебр.

Означення (вектора з незалежними компонентами)

Вектор $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ називається вектором з незалежними компонентами, якщо $(\xi_j)_{j=1}^n$ — послідовність незалежних випадкових величин.

Вправи.

1. Нехай $g\colon \mathbb{R}^m \to \mathbb{R}$ — борелева функція, а $\xi = (\xi_1, \ldots, \xi_m)$ — випадковий вектор. Тоді безпосередньо з означень випливає, що $\eta \stackrel{def}{=} g(\xi)$ — випадкова величина. Довести.

2. Нехай $g\colon \mathbb{R}^m \to \mathbb{R}^n$ — борелева функція, а $\xi = (\xi_1, \ldots, \xi_m)$ — випадковий вектор. Чи вірно, що тоді $\eta \stackrel{def}{=} g(\xi)$ — випадковий вектор?

3. Нехай (η_j) — послідовність випадкових величин. Довести, що функція $\eta\colon \Omega \to \mathbb{R}$, визначена однією з наступних рівностей: (a) $\eta(\omega) = \sum_{j=1}^{n} \eta_j(\omega)$, (b) $\eta(\omega) = \prod_{j=1}^{n} \eta_j(\omega)$, (c) $\eta(\omega) = \lim\limits_{j\to+\infty} \eta_j(\omega)$, (d) $\eta(\omega) = \sum_{j=1}^{+\infty} \eta_j(\omega)$, (e) $\eta(\omega) = \overline{\lim\limits_{j\to+\infty}}\, \eta_j(\omega)$, (f) $\eta(\omega) = \underline{\lim\limits_{j\to+\infty}}\, \eta_j(\omega)$, (g) $\eta(\omega) = \sup\{\eta_j(\omega)\colon j \in \mathbb{N}\}$, (h) $\eta(\omega) = \inf\{\eta_j(\omega)\colon j \in \mathbb{N}\}$, — є випадковою величиною.

4. Якщо $(\mathcal{A}_j)_{j=1}^{n}$ — послідовність незалежних σ-алгебр і $(\forall j,\ 1 \le j \le n)\colon\ \widetilde{\mathcal{A}}_j \subset \mathcal{A}_j$, $\widetilde{\mathcal{A}}_j$ — σ-алгебри, то $(\widetilde{\mathcal{A}}_j)_{j=1}^{n}$ — послідовність незалежних σ-алгебр.

5. Для випадкової величини ξ позначимо $\xi^+ = \sup\{0, \xi\}$, $\xi^- = \sup\{0, -\xi\}$, Якщо (ξ_j) — послідовність незалежних випадкових величин, то $(\xi_j^{\pm})$ — також послідовність незалежних випадкових величин для будь-якого фіксованого вибору послідовності плюсів і мінусів у верхньому індексі. Довести.

2.2. Функції розподілу випадкових величин та їхні властивості

Оскільки, множина $(-\infty, x)$, $x \in \mathbb{R}$ є борелевою, то можна означити поняття функції розподілу випадкової величини.

Означення (функції розподілу випадкової величини)

Функцією розподілу випадкової величини ξ називається функція $F_\xi : \Omega \to \mathbb{R}$, визначена для кожного $x \in \mathbb{R}$ формулою

$$F_\xi(x) \stackrel{def}{=} \mathbb{P}\{\omega : \xi(\omega) < x\}.$$

Теорема (про характеристичні властивості)

Нехай ξ — випадкова величина. Функція розподілу $F_\xi(x)$ має такі властивості: **1**. $F_\xi(+\infty) = 1$; **2**. $F(-\infty) = 0$; **3**. $F_\xi(x)$ — неспадна;
4. $F_\xi(x)$ — неперервна зліва тобто $(\forall x_0)$: $\lim\limits_{x \to x_0 - 0} F_\xi(x) = F_\xi(x_0)$.

Доведення Доведемо спочатку властивість 3., позаяк з неї, зокрема, випливає, що граничні значення функції розподілу з інших трьох властивостей існують.
Нехай $a < b$. Тоді $F_\xi(b) - F_\xi(a) = \mathbb{P}\{\omega\colon \xi(\omega) < b\} - \mathbb{P}\{\omega\colon \xi(\omega) < a\} = \mathbb{P}\{\omega\colon a \le \xi(\omega) < b\} \ge 0$. Отже, $F_\xi(x)$ — неспадна.
Доведемо властивість 1. Досить довести, що $\forall (x_n)$ монотонно зростаючої до $+\infty$: $F_\xi(x_n) \to 1$ при $n \to +\infty$.
Розглянемо послідовність множин $A_n = \{\omega : \xi(\omega) < x_n\}$; $A_n \subset A_{n+1}$ $(n \ge 1)$, тобто, (A_n) — монотонно зростаюча послідовність. Оскільки $\bigcup\limits_{n=1}^{+\infty} A_n = \xi^{-1}(\mathbb{R}) = \Omega$, то за неперервністю ймовірності

$$\lim_{n\to+\infty} F_\xi(x_n) = \lim_{n\to+\infty} \mathbb{P}(A_n) = \mathbb{P}(\bigcup_{n=1}^{+\infty} A_n) = \mathbb{P}(\Omega) = 1.$$

Для того, щоб довести властивість 2., досить довести, що

$$(\forall (y_n), y_n \downarrow -\infty \ (n \to +\infty)) : F_\xi(-\infty) \stackrel{def}{=} \lim_{n\to+\infty} F_\xi(y_n) = 0.$$

Розглянемо монотонно спадну послідовність множин $B_n = \{\omega : \xi(\omega) < y_n\}$, $B_{n+1} \subset B_n \ (n \geq 1)$. Оскільки, $\bigcap\limits_{n=1}^{+\infty} B_n = \varnothing$, то за неперервністю ймовірності маємо

$$\lim_{n\to+\infty} F_\xi(y_n) = \lim_{n\to+\infty} \mathbb{P}(B_n) = \mathbb{P}(\bigcap_{n=1}^{+\infty} B_n) = \mathbb{P}(\varnothing) = 0.$$

Для того, щоб довести властивість 4., досить довести, що

$$(\forall x_0)(\forall (z_n), z_n \uparrow x_0 \ (n \to +\infty)) : \lim_{n\to+\infty} F_\xi(z_n) = F_\xi(x_0).$$

Знову розглянемо монотонно зростаючу послідовність множин $C_n = \{\omega : \xi(\omega) < z_n\}$. Оскільки $\bigcup\limits_{n=1}^{+\infty} C_n = \{\omega : \xi(\omega) < x_0\}$,

то за неперервністю ймовірності отримуємо

$$F_\xi(x_0) = \mathbb{P}\{\omega : \xi(\omega) < x_0\} = \mathbb{P}(\bigcup_{n=1}^{+\infty} C_n) = \lim_{n\to+\infty} \mathbb{P}(C_n) = \lim_{n\to+\infty} F_\xi(z_n).$$

Властивості 1. — 4. з теореми 2 ще називають характеристичними властивостями функцій розподілу, позаяк вони виділяють з класу всіх дійсних функцій, заданих на всій числовій прямій, функції, кожна з яких є функцією розподілу для деякої випадкової величини. Власне, справджується така обернена теорема.

Теорема (обернена до теореми про характеристичні властивості)

Нехай $F(x)$ — функція з $\mathbb{R} \to \mathbb{R}_+$, яка має властивості 1.–4. з теореми про характеристичні властивості. Тоді існує ймовірнісний простір $(\Omega, \mathcal{A}, \mathbb{P})$ і існує ξ — випадкова величина на ньому така, що $F_\xi(x) \equiv F(x)$.

Доведення. Нехай $\mathbb{P} = \nu_F$ міра Лебега-Стілт'єса, породжена функцією F. Міра Лебега-Стільтєса ν_F на σ-алгебрі борелевих множин на прямій ν_F запроваджується за такою схемою. Міра півінтервалу $\nu_F([a, b)) \stackrel{def}{=} F(b) - F(a)$.
На $\mathcal{A}_0$ — алгебрі, породженій скінченними об'єднаннями півінтервалів вводиться природно, наприклад,
$\nu_F\left(\bigsqcup_{k=1}^{n}[a_k, b_k)\right) \stackrel{def}{=} \sum_{k=1}^{n} \nu_F([a_k, b_k))$.
Далі, перевіряється, що ν_F — зліченно-адитивна ймовірнісна міра на $\mathcal{A}_0$.
І, нарешті, тепер вже можна скористатись теоремою Каратеодорі-Лебега, яку ми сформулюємо у такому вигляді.

Теорема (Каратеодорі-Лебега про продовження міри)

Якщо Q — ймовірнісна зліченно-адитивна міра на алгебрі $\mathcal{A}_0$, то існує єдина $\mathbb{P}$ — ймовірнісна зліченно-адитивна міра на σA_0 (σA_0 — мінімальна σ-алгебра, породжена алгеброю A_0) така, що

$$(\forall A \in \mathcal{A}_0)\colon\ \mathbb{P}(A) = Q(A).$$

Тоді, за теоремою Каратеодорі-Лебега про продовження міри, існує єдине продовження $\widetilde{\nu}_F$ міри ν_F на мінімальну σ-алгебру, породжену σ-алгеброю $\mathcal{A}_0$, тобто існує єдине продовження ν_F на σ-алгебру борелевих множин $\mathcal{B}(\mathbb{R})$ на прямій таке, що

$$(\forall A \in \mathcal{A}_0)\colon\ \widetilde{\nu}_F(A) = \nu_F(A).$$

Скрізь нижче у позначенні $\widetilde{\nu}_F$ хвильку пропускатимемо. Отже, за побудовою $\mathbb{P} = \nu_F$ ймовірність, визначена на σ-алгебрі борелевих множин на прямій.
$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \nu_F)$ розглянемо тепер випадкову величину $\xi(\omega) = \omega$. Тоді,

$$F_\xi(x) = \mathbb{P}\{\omega\colon \omega < x\} = \nu_F((-\infty, x)) = F(x) - F(-\infty) = F(x).$$

Вправа. Довести зліченну адитивність ν_F на $\mathcal{A}_0$, використавши теорему про неперервність зліченно-адитивної ймовірнісної міри.

Вважають, що випадкову величину описано, якщо задано її функцію розподілу.

2.3. Властивості функції розподілу (продовження)

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

Доведення цього факту знаходимо у доведенні властивості **3**. з теореми про характеристичні властиворсті функції розподілу. Справді, $F_\xi(b) - F_\xi(a) = \mathbb{P}\{\omega\colon \xi(\omega) < b\} - \mathbb{P}\{\omega\colon \xi(\omega) < a\} = \mathbb{P}\Big(\{\omega\colon \xi(\omega) < b\} \setminus \{\omega\colon \xi(\omega) < a\}\Big) = \mathbb{P}\{\omega\colon a \le \xi(\omega) < b\}$.

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

Для доведення досить для довільної послідовності (y_n) такої, що $y_n \uparrow b$ $(n \to +\infty)$ розглянути монотонно спадну послідовність множин $B_n = \{\omega\colon b \le \xi(\omega) < y_n\}$. Оскільки, $B \stackrel{def}{=} \bigcap\limits_{n=1}^{+\infty} B_n = \{\omega\colon \xi(\omega) = b\}$, то використовуючи властивість 5. і знову теорему неперервності ймовірності, отримуємо

$$\mathbb{P}(\{\omega\colon \xi(\omega) = b\}) = \mathbb{P}(B) = \mathbb{P}\Big(\bigcap_{n=1}^{+\infty} B_n\Big) =$$

$$= \lim_{n\to+\infty} \mathbb{P}(B_n) = \lim_{n\to+\infty} (F_\xi(y_n) - F_\xi(b)) = F_\xi(b+0) - F_\xi(b).$$

З властивості **6**. негайно отримуємо таке твердження.

7. Якщо b – точка неперервності F_ξ, то $\mathbb{P}\{\omega : \xi(\omega) = b\} = 0$.

Справді, $\mathbb{P}\{\omega : \xi(\omega) = b\} = F_\xi(b+0) - F_\xi(b) = 0$.

З властивостей **5**. і **6**. негайно отримуємо таке твердження.

8. $\mathbb{P}\{\omega : a \le \xi(\omega) \le b\} = F_\xi(b+0) - F_\xi(a)$.

Справді, $\mathbb{P}\{\omega : a \le \xi(\omega) \le b\} = \mathbb{P}\{\omega : a \le \xi(\omega) < b\} + \mathbb{P}\{\omega : \xi(\omega) = b\} = F_\xi(b) - F_\xi(a) + F_\xi(b+0) - F_\xi(b) = F_\xi(b+0) - F_\xi(a)$.

Вправа. Якщо $\omega = \omega_0$ — точка неперервності випадкової величини ξ, то чи твердження **7**. залишиться правильним при $b = \xi(\omega_0)$? Тобто, чи правильно, що $\mathbb{P}\{\omega : \xi(\omega) = \xi(\omega_0)\} = 0$ у випадку, коли $\omega = \omega_0$ — точка неперервності випадкової величини ξ?

Example (рівномірний розподіл)

Розглянемо ймовірнісний простір $\Omega = \mathbb{R}$, $\mathbb{P}$ – міра Лебега (довжина). Нехай випадкова величина $\xi(\omega)\colon [a, b] \to \mathbb{R}$ така, що $\xi(\omega) = \omega$. Тоді, $F_\xi(x) = \mathbb{P}\{\omega\colon \xi(\omega) < x\} =$

$$= \begin{cases} \mathbb{P}(\varnothing) = 0, & x \le a, \\ \mathbb{P}([a, x)) = \dfrac{x - a}{b - a}, & a < x \le b, \\ \mathbb{P}([a, b)) = \dfrac{b - a}{b - a} = 1, & x > b. \end{cases}$$

2.5. Дискретні випадкові величини

Нехай $(\Omega, \mathcal{A}, \mathbb{P})$ — фіксований ймовірнісний простір.

Означення (дискретної випадкової величини)

Випадкову величину $\xi(\omega)$ називатимемо дискретною, якщо її множина значень$X = \left\{x_k\colon\ (\exists\omega)[\xi(\omega) = x_k]\right\}$ не більш, ніж зліченна.

Твердження (критерій дискретної випадкової величини)

Функція $\xi = \xi(\omega)\colon \Omega \to X$ з не більш, ніж зліченною множиною значень $X = \{x_n\colon 1 \le n < N\}$ $(N \le +\infty)$ є дискретною випадковою величиною тоді і тільки тоді, коли $(\forall n,\ 1 \le n < N)\colon\ \{\omega : \xi(\omega) = x_n\} \in \mathcal{A}$.

Доведення. Необхідність. Нехай $n \in \mathbb{N},\ n < N,$ — довільне фіксоване число. Оскільки одноточкова множина $\{x_n\} \in \mathcal{B}$ — борелева, то за означенням випадкової величини $\xi^{-1}(\{x_n\}) \in \mathcal{A}$.

Достатність. За критерієм випадкової величини досить перевірити, що $(\forall x \in \mathbb{R})\ \{\omega : \xi(\omega) < x\} \in \mathcal{A}$. Але $\{\omega : \xi(\omega) < x\} = \bigcup\limits_{x_n < x} \{\omega : \xi(\omega) = x_n\}$, а кожна така множина $\{\omega : \xi(\omega) = x_n\} \in \mathcal{A}$. За означенням σ-алгебри, не більш, ніж зліченні об'єднання множин з σ-алгебри належать до σ-алгебри. Отже, $\{\omega : \xi(\omega) < x\} \in \mathcal{A}$.

Нехай ξ — дискретна випадкова величина, а $\{x_n : n \geq 1\}$ — множина її значень. Позначимо $p_k = \mathbb{P}\{\omega : \xi(\omega) = x_k\}$. Тоді, враховуючи попарну несумісність подій $(\{\omega : \xi(\omega) = x_n\})$ та зліченну адитивність ймовірності, отримаємо

$$(\forall x \in \mathbb{R}):\ \ F_\xi(x) = \mathbb{P}\{\omega : \xi(\omega) < x\} =$$

$$= \mathbb{P}\Big(\bigcup_{x_k < x} \{\omega : \xi(\omega) = x_k\}\Big) = \sum\nolimits_{x_k < x} p_k.$$

Вправа. Довести, що випадкова величина ξ є дискретною випадковою величиною тоді і тільки тоді, коли існує скінченна або зліченна послідовність (x_n) така, що її функція розподілу має вигляд

$$F_\xi(x) = \sum\nolimits_{x_k<x} p_k,$$

при $p_k = \mathbb{P}(\{\omega : \xi(\omega) = x_k\})$, $\sum_{k\geq 1} p_k = 1$.

Дискретна випадкова величина часто записується табличкою: задається послідовність значень випадкової величини і відповідна послідовність значень її ймовірностей.

Зауважимо, що $F_\xi(x+0) - F_\xi(x) = \begin{cases} p_k, \text{ якщо } x = x_k, \\ 0, \ x \notin \{x_k\}. \end{cases}$

У випадку, коли (x_k) — монотонно зростаюча послідовність, то $F_\xi(x)$ — функція стрибків (кусково стала), тобто

$$F_\xi(x) = \sum\nolimits_{k=1}^{n} p_k, \quad \text{для } \ x \in [x_n, x_{n+1}).$$

Означення (простої випадкової величини)

Випадкову величину $\xi(\omega)$ називатимемо простою, якщо її множина значень $X = \left\{x_k\colon\ (\exists\omega)[\xi(\omega) = x_k]\right\}$ містить скінченну кількість елементів.

Кожна проста випадкова величина є дискретною.
Нехай ξ – проста випадкова величина на ймовірнісному просторі $(\Omega, \mathcal{A}, \mathbb{P})$, а $X = \{x_n\colon 1 \le n \le N\},\ N < +\infty$ – її множина значень; $p_n = \mathbb{P}\{\xi(\omega) = x_n\}$.
Нехай також $g\colon \mathbb{R} \to \mathbb{R}$ – борелева функція. Тоді, за означенням

$$\int\limits_{\Omega} g(\xi(\omega))\mathbb{P}(d\omega) \stackrel{def}{=} \sum_{n=1}^{N} g(x_n)\cdot p_n = \sum_{n=1}^{N} g(x_n)\cdot \mathbb{P}\{\xi(\omega) = x_n\}.$$

Відзначимо також, що для дискретної випадкової величини, якщо $A \in \mathcal{A}$, то за означенням інтеграла

$$\int_A g(\xi(\omega))\mathbb{P}(d\omega) = \int_\Omega \chi_A(\omega) g(\xi(\omega))\mathbb{P}(d\omega) = \sum\nolimits_{x_n \in \xi(A)} g(x_n)\cdot p_n.$$

2.4. Інтеграли, що зустрічаються в теорії ймовірностей

Розглянемо тепер означення деяких інтегралів, які використовуються в теорії ймовірностей.
Якщо $\xi(\omega)$ — випадкова величина на ймовірнісному просторі $(\Omega, \mathcal{A}, \mathbb{P})$, а $g(x)$ — борелева функція на дійсній прямій, то

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

— абстрактний інтеграл Лебега за ймовірнісною мірою $\mathbb{P}$,

$$\int_{-\infty}^{+\infty} g(x)dF(x) \stackrel{def}{=} \int_{-\infty}^{+\infty} g(x)\nu_F(dx)$$

— інтеграл Лебега-Стілт'єса, тобто, абстрактний інтеграл Лебега за ймовірнісною мірою Лебега-Стілт'єса ν_F, породженою функцією розподілу F. За формулою заміни змінної інтегрування

$$\int_{\Omega} g(\xi(\omega))\mathbb{P}(d(\omega)) = \int_{-\infty}^{+\infty} g(x)dF_{\xi}(x).$$

1. Якщо $F(x)$ — неперервно диференційовна на $\mathbb{R}$ за винятком послідовності точок (c_n), то
$\int_{-\infty}^{+\infty} g(x)dF(x) = \int_{-\infty}^{+\infty} g(x)F'(x)dx + \sum\limits_n g(c_n)(F(c_n+0) - F(c_n))$.

2. $F(x)$ — кусково-стала функція, для якої (c_n) — послідовність точок стрибка. Тоді

$$\int_{-\infty}^{+\infty} g(x)dF(x) = \sum_n g(c_n)(F(c_n+0) - F(c_n)).$$

3. Нехай $A \in \mathcal{A}$ і χ_A — її характеристична функція. Тоді

$$\int_A g(\xi(\omega))\mathbb{P}(d\omega) \stackrel{def}{=} \int_\Omega g(\xi(\omega))\chi_A(\omega)\mathbb{P}(d\omega).$$

Отже перейдемо до означення записаних інтегралів.

Абстрактний інтеграл Лебега

I. Нехай $\eta(\omega)$ – невід'ємна випадкова величина, тобто, $\eta(\omega) \geq 0$. Тоді існує монотонно неспадна послідовність простих невід'ємних випадкових величин $(\eta_n(\omega))$ таких, що $\lim\limits_{n\to+\infty} \eta_n(\omega) = \eta(\omega)$. Власне,

1) $(\forall n)(\forall \omega)\colon \eta_n(\omega) \geq 0$,

2) $(\forall n)\colon \eta_n(\omega)$ – проста випадкова величина,

3) $(\forall \omega)\colon\ \eta_n(\omega) \uparrow \eta(\omega)\ (n \to +\infty)$.

Інтеграли $I_n := \int\limits_{\Omega} \eta_n(\omega) d\omega$ вже визначені, бо $\eta_n(\omega)$ – проста випадкова величина, і ці інтеграли є фактично сумами невід'ємних величин, а тому $I_n \geq 0$. З іншого боку послідовність $(\eta_n(\omega))$ є неспадна. Тому числова послідовність (I_n) також неспадна.

Вправа. Довести цей факт!

Тоді, у випадку, якщо числова послідовність (I_n) є обмеженою, то існує скінченна границя $\lim\limits_{n\to+\infty} I_n = I \in [0, +\infty)$, а якщо ця послідовність необмежена, то границя існує і дорівнює $I = +\infty$. в обидвох випадках покладаємо

$$\int\limits_{\Omega} \eta(\omega)\mathbb{P}(d\omega) := I.$$

II. Нехай $\eta(\omega)$ – довільна випадкова величина.
Означимо $\eta^+(\omega) = \sup\{0, \eta(\omega)\}$, $\eta^-(\omega) := (-\eta(\omega))^+$ або інакше $\eta^-(\omega) := \sup\{0, -\eta(\omega)\}$. Тоді, $\eta^{\pm}(\omega) \geq 0$ і за I. існують інтеграли $I^+ = \int\limits_{\Omega} \eta^+(\omega)\mathbb{P}(d\omega)$, $I^- = \int\limits_{\Omega} \eta^-(\omega)\mathbb{P}(d\omega)$.
У випадку, коли один з інтегралів $I^{\pm} < +\infty$ за означенням приймемо

$$\int\limits_{\Omega} \eta(\omega)\mathbb{P}(d\omega) := I^+ - I^-.$$

А також, для довільної $A \in \mathcal{A}$ приймемо

$$\int\limits_{A} \eta(\omega)\mathbb{P}(d\omega) := \int\limits_{\Omega} \chi_A(\omega)\eta(\omega)\mathbb{P}(d\omega).$$

III. Властивості інтеграла.

1. Якщо випадкова величина $\eta(\omega) \geq 0$, то інтеграл визначений (тобто, випадкова величина інтегровна) і $\int\limits_{\Omega} \eta(\omega)\mathbb{P}(d\omega) \geq 0$.

2. Нехай для випадкових величин η_1, η_2 їхні інтеграли є скінченні, а $\alpha, \beta \in \mathbb{R}$. Тоді,

$$\int\limits_{\Omega} (\alpha\eta_1(\omega) + \beta\eta_2(\omega))\mathbb{P}(d\omega) = \alpha \int\limits_{\Omega} \eta_1(\omega)\mathbb{P}(d\omega) + \beta \int\limits_{\Omega} \eta_2(\omega)\mathbb{P}(d\omega).$$

Оскільки ці дві властивості для інтегралів від простих випадкових величин (нагадаємо, що вони є скінченними сумами) очевидно, що виконуються, а інтеграли за означенням отримуються з таких самих рівностей для простих випадкових величин граничним переходом, а границі числових послідовностей мають такі самі властивості, то на цьому шляху отримаємо ці дві властивості.

Вправа. Довести ці дві властивості.

З цих двох властивостей випливає така властивість монотонності інтеграла.

3. $(\forall\omega)\colon \eta_1(\omega) \le \eta_2(\omega) \Longrightarrow \int\limits_{\Omega} \eta_1(\omega)\mathbb{P}(d\omega) \le \int\limits_{\Omega} \eta_2(\omega)\mathbb{P}(d\omega)$.

Справді, досить розглянути випадкову величину $\eta = \eta_2 - \eta_1 \ge 0$. Тоді, послідовно маємо $0 \le \int_\Omega \eta\mathbb{P}(d\omega) = \int_\Omega \eta_2\mathbb{P}(d\omega) - \int_\Omega \eta_1\mathbb{P}(d\omega) \ge 0$.

4. Нехай $A_1, A_2 \in \mathcal{A},\ A_1 \cap A_2 = \varnothing \Longrightarrow$

$$\int\limits_{A_1\cup A_2} \eta\mathbb{P}(d\omega) = \int\limits_{A_1} \eta\mathbb{P}(d\omega) + \int\limits_{A_2} \eta\mathbb{P}(d\omega).$$

Справді, $\chi_{A_1\cup A_2} = \chi_{A_1} + \chi_{A_2}$, тому

$$\int\limits_{A_1\cup A_2} \eta\mathbb{P}(d\omega) = \int_\Omega \chi_{A_1\cup A_2}\eta\mathbb{P}(d\omega) = \int_\Omega(\chi_{A_1} + \chi_{A_2})\eta\mathbb{P}(d\omega) =$$

$$= \int_\Omega \chi_{A_1}\eta\mathbb{P}(d\omega) + \int_\Omega \chi_{A_2}\eta\mathbb{P}(d\omega) = \int\limits_{A_1} \eta\mathbb{P}(d\omega) + \int\limits_{A_2} \eta\mathbb{P}(d\omega).$$

5. $\left|\int_\Omega \eta \mathbb{P}(d\omega)\right| \leq \int_\Omega |\eta| \mathbb{P}(d\omega)$. При цьому, якщо існує інтеграл від $|\eta|$, то існує інтеграл від η.
Справді, $\eta = \eta^+ - \eta^-$, $|\eta| = \eta^+ + \eta^-$. Далі, зрозуміло, що за означенням, інтеграл від $|\eta|$ існує у випадку, коли існують інтеграли від η^+ і η^-, бо $\eta^\pm \geq 0$. Але тоді за означенням існує інтеграл від η.
6. $\int_A \mathbb{P}(d\omega) = \mathbb{P}(A)$.
За означенням $\displaystyle\int_A \mathbb{P}(d\omega) = \int_\Omega \chi_A \mathbb{P}(d\omega)$. Але, χ_A – проста функція, що набуває двох значень $0, 1$. Тоді, за означенням $\int_\Omega \chi_A \mathbb{P}(d\omega) = 1 \cdot \mathbb{P}(A) + 0 \cdot \mathbb{P}(\overline{A}) = \mathbb{P}(A)$.

Інтеграл Лебега-Стілт'єса:

$$\int_{-\infty}^{+\infty} g(x) dF(x) \stackrel{def}{=} \int_{-\infty}^{+\infty} g(x) \nu_F(dx).$$

Вище ми побудували міру Лебега-Стілт'єса ν_F на σ-алгебрі $\mathcal{B}(\mathbb{R})$ борелевих множин на прямій, породжену функцією розподілу F. Вона виявилася ймовірністю. Тоді, інтеграл $\int_{-\infty}^{+\infty} g(x)\nu_F(dx)$ є абстрактним інтегралом від $g(x)$ на прямій $\mathbb{R}$ за ймовірністю ν_F.

Інтеграл Лебега на прямій $\mathbb{R}$.

Означимо інтеграл $\int\limits_a^b g(x)dx = \int\limits_{[a,b]} g(x)dx$. Означимо спочатку інтеграл $\int\limits_0^1 g(x)dx = \int\limits_{[0,1]} g(x)dx$.Для цього розглянемо функцію розподілу $F(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 1, \\ x, & 0 < x \leq 1 \end{cases}$ та міру Лебега-Стілт'єса, породжену нею. Міра Лебега-Стілт'єса, породжена нею називається мірою Лебега $\mu_{[0,1]}$, визначеною на σ-алгебрі борелевих множин на проміжку $[0,1]$.

За означенням, для $[a,b] \subset [0,1]$ маємо

$$\int\limits_a^b g(x)dx = \int\limits_{[a,b]} g(x)dx = \int\limits_0^1 \chi_{[a,b]} g(x)\nu_F(dx).$$

Зокрема, $\int\limits_a^b dx = \int\limits_0^1 \chi_{[a,b]}\nu_F(dx) = F(b) - F(a) = b - a.$

Для довільної борелевої множини $B \subset [a, b]$ розглянемо її образ B^* при відображенні $x = x(t) = \dfrac{t-a}{b-a} \colon B \to B^* \subset [0, 1]$. При цьому $x([a, b]) = [0, 1]$. Тоді, мірою Лебега
$\mu(B) := (b - a)\mu_{[0,1]}(B^*)$.
Відзначимо, що у випадку, коли функція $g(x)$ інтегровна за Ріманом, то її інтеграл Лебега дорівнює її інтегралу Рімана.

2.6. Неперервні випадкові величини: абсолютно неперервні та сингулярні

Означення (абсолютно неперервної випадкової величини)

Випадкова величина ξ називається абсолютно неперервною, або випадковою величиною зі щільністю, якщо існує така невід'ємна вимірна (за Лебегом) функція $f_\xi \colon \mathbb{R} \to \mathbb{R}_+$ така, що $(\forall B \in \mathcal{B}(\mathbb{R}))$:

$$\mathbb{P}\{\omega \colon \xi(\omega) \in B\} = \int_B f_\xi(t)dt.$$

Функцію $f_\xi(x)$ називатимемо щільністю або густиною розподілу. Також у цьому контексті вживатимемо термін абсолютно неперервний розподіл.

Безпосередньо з означення випливають такі властивості щільності розподілу:

1. $F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)dt$ (досить вибрати в означенні $B = (-\infty, x)$),

звідки, зокрема, маємо, що

2. $F'_\xi(x) = f_\xi(x)$ м.с. (майже скрізь за мірою Лебега),

3. $\int_{-\infty}^{+\infty} f_\xi(t)dt = 1$.

Теорема (Лебега)

Довільну функцію розподілу F однозначно можна зобразити у вигляді

$$F(x) = \sum_{j=1}^{3} a_j F_j(x),$$

де сталі $a_1, a_2, a_3 \geq 0$ такі, що $a_1 + a_2 + a_3 = 1$, а F_1 — дискретний розподіл (функція стрибків), F_2 — абсолютно неперервний розподіл, F_3 — сингулярна функція розподілу (тобто, неперервна функція розподілу без щільності).

Вправи. 1. Нехай ξ випадкова величина, а $h\colon \mathbb{R} \to \mathbb{R}_+$ вимірна за Лебегом функція така, що

$$(\forall x \in \mathbb{R})\colon\ F_\xi(x) = \int_{-\infty}^{x} h(t)dt.$$

Довести, що випадкова величина ξ має щільність і $h(t) = f_\xi(t)$ м.с.

2. Нехай ξ — абсолютно неперервна випадкова величина. Чи $(\forall x \in \mathbb{R})\colon\ \mathbb{P}\{\omega : \xi(\omega) = x\} = 0$?

3. Нехай $\mathcal{K}\colon [0,1] \to [0,1]$ — канторова "драбина". Визначимо

$F(x) = 0$ для $x < 0$, $F(x) = \mathcal{K}(x)$ для $x \in [0, 1]$, $F(x) = 1$ для $x > 1$. Довести, що розподіл F — сингулярний.

4. Для абсолютно неперервного розподілу F_ξ і борелевої функції g

$$\int_\Omega g(\xi(\omega))\mathbb{P}(d\omega) = \int_{-\infty}^{+\infty} g(x) f_\xi(x) dx.$$

Вказівка. Для доведення цієї рівності досить переконатись, що $dF_\xi(x) = f_\xi(x)dx$, тобто, що $\int_B dF_\xi(x) = \int_B f_\xi(x)dx$ для кожної борелевої множини B.

2.7. Рівномірний розподіл

На відрізку $[a, b]$ на прямій навмання вибирають точку. Вважаючи вибір будь-якої точки на відрізку однаково можливим, розглядаємо ймовірнісний простір $([a, b], \mathcal{A}, \mathbb{P})$, де $\mathcal{A}$ — σ-алгебра вимірних за Лебегом підмножин відрізка, а за формулою геометричних ймовірностей $\mathbb{P}(A) = \dfrac{\mathrm{meas}_1 A}{b - a}$.

Розглянемо випадкову величину $\xi(\omega) = \omega$, $\omega \in [a, b]$. Тоді її

функція розподілу

$$F_\xi(x) = \mathbb{P}\{\omega : \xi(\omega) < x\} = \begin{cases} \frac{\mathrm{meas}_1(\varnothing)}{b-a} = 0, & \text{якщо } x \le a \\ \frac{\mathrm{meas}_1([a,x))}{b-a} = \frac{x-a}{b-a}, & \text{якщо } a < x \le b \\ \frac{\mathrm{meas}_1([a,b])}{b-a} = 1, & \text{якщо } x > b. \end{cases}$$

Даний розподіл називається рівномірним розподілом на відрізку $[a, b]$. Очевидно, що рівномірний розподіл має щільність

$$f_\xi(x) = F'_\xi(x) = \begin{cases} 0, & \text{якщо } x < a \\ \frac{1}{b-a}, & \text{якщо } a < x < b \\ 0, & \text{якщо } x > b. \end{cases}$$

2.8. Розподіл Коші

В точці на площині з координатами (a, b), $b > 0$, розташоване джерело радіоактивних частинок, з якого у напрямку осі Ox (горизонтальна мішень) вилітають α-частинки, які рухаються прямолінійно. Нехай $\xi(\omega)$ — означає випадкову абсцису точки мішені, у яку потрапляє α-частинка. Припустимо, що випадкові значення кутів $\varphi = \varphi(\omega)$, які утворюють напрямки руху α-частинок з напрямком "вниз" на вертикальній прямій, що проходить через точку (a, b), — випадкова величина з рівномірним розподілом на $\left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$. Тоді, оскільки щільність $f_\varphi(t) = \frac{1}{\pi}$, для $-\frac{\pi}{2} < t < \frac{\pi}{2}$, $f_\varphi(t) = 0$, для $t \notin [-\frac{\pi}{2}; \frac{\pi}{2}]$, і

$$(\forall u \in \mathbb{R})\colon\ F_\xi(u) = \mathbb{P}\{\omega\colon \xi(\omega) < u\} = \mathbb{P}\{\omega\colon a + b\,\mathrm{tg}\,\varphi(\omega) < u\} =$$

$$= \mathbb{P}\left\{\omega\colon \varphi(\omega) < \mathrm{arctg}\,\frac{u-a}{b}\right\} = F_\varphi\Big(\mathrm{arctg}\,\frac{u-a}{b}\Big),$$

то щільність розподілу випадкової величини ξ

$$(\forall u \in \mathbb{R})\colon f_\xi(u) = F'_\xi(u) =$$

$$= F'_\varphi\Big(\operatorname{arctg}\frac{u-a}{b}\Big)\frac{1/b}{1+\left(\frac{u-a}{b}\right)^2} = \frac{b}{\pi}\cdot\frac{1}{b^2+(u-a)^2}.$$

Розподіл з даною щільністю називається

розподілом Коші $\mathcal{K}(a,b)$;

$\mathcal{K}(0,1)$ — стандартний розподіл Коші зі щільністю:

$$f(x) = \frac{1}{\pi}\cdot\frac{1}{1+x^2}.$$

2.9. Схема Бернуллі

Здійснюється послідовність з $n \in \mathbb{N}$ послідовних незалежних експериментів (випробувань), у кожному з яких може появитися подія A з ймовірністю $\mathbb{P}(A) = p = p(n)$, та не появитися з ймовірністю $\mathbb{P}(\overline{A}) = q = 1 - p$. При цьому появу події A стандартно вважають успіхом (позначення — У), а появу $\overline{A}$ вважають неуспіхом (позначення — Н).
З кожним з випробувань пов'язуємо свою σ-алгебру $\mathcal{A}_{j,n}$. За означенням незалежності експериментів $(\mathcal{A}_{j,n})_{j=1}^{n}$ — послідовність незалежних σ-алгебр.
Це класична схема Бернуллі.
Нехай m_n — кількість У в $n \in \mathbb{N}$ послідовних випробуваннях.
Введемо у розгляд випадкові величини $\xi_{j,n}$ — індикатори появи події $A \in \mathcal{A}_{j,n}$, тобто, в j-тому випробуванні у серії з n випробувань.

Оскільки $\mathcal{A}_{\xi_{j,n}} \subset \mathcal{A}_{j,n}$, то $(\xi_{j,n})_{j=1}^{n}$ — послідовність незалежних випадкових величин. Крім цього зауважимо, що оскільки

$$m_n = \sum\nolimits_{j=1}^{n} \xi_{j,n},$$

то m_n — дискретна випадкова величина з множиною значень $\{0, 1, \ldots, n\}$.

1. Біномний розподіл. Знайдемо розподіл випадкової величини m_n. Для цього потрібно відшукати ймовірності

$$\mathbb{P}_n(k) \stackrel{def}{=} \mathbb{P}\{\omega\colon m_n(\omega) = k\} \quad (k \in \{0, 1, \ldots, n\}).$$

З незалежності послідовних випробувань випливає, що ймовірність появи точно k У в n послідовних випробуваннях не залежить від того, у яких з послідовних випробувань У появиться. Опишемо цей факт точніше. Через $A_{j,n}$ позначимо подію A, як елемент σ-алгебри $\mathcal{A}_{j,n}$. А через U, V позначимо доповняльні множини індексів, тобто
$U \cup V = \{1, 2, \ldots, n\}$, $U \cap V = \varnothing$, $|U| = k$.

Тоді, з незалежності випробувань випливає, що для будь-яких таких множин індексів

$$\mathbb{P}\Big(\bigcap_{j\in U} A_{j,n} \bigcap_{i\in V} \overline{A}_{i,n}\Big) = p^k q^{n-k}.$$

Залишилось пригадати, що число різних k-елементних піджмножин $U \subset \{1, 2, \ldots, n\}$ дорівнює біномному коефіцієнту C_n^k. Отже,

$$\mathbb{P}_n(k) = C_n^k p^k q^{n-k} \quad (k \in \{0, 1, \ldots, n\}).$$

Зауважимо, що

$$\sum\nolimits_{k=0}^{n} \mathbb{P}_n(k) = \sum\nolimits_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1.$$

Власне, тому розподіл випадкової величини m_n і називається біномним розподілом.

2. Геометричний розподіл. Нехай ξ — кількість неуспіхів (Н) до першої появи успіхів (У). Отримуємо добуток послідовних результатів експерименту, який символічно записуємо у вигляді $\underbrace{\text{Н...Н}}_{k}$У. З незалежності випробувань, як і вище отримуємо,

$$\mathbb{P}\{\omega\colon \xi(\omega) = k\} = \mathbb{P}\{\underbrace{\text{Н...Н}}_{k}\text{У}\} = q^k p \quad (k \in \mathbb{Z}_+).$$

Це так званий геометричний розподіл.

Вправа. Знайти k_0 таке, що $\max\{\mathbb{P}_n(k)\colon 0 \le k \le n\} = \mathbb{P}_n(k_0)$.

Вказівка. Розглянути частку $\dfrac{\mathbb{P}_n(k)}{\mathbb{P}_n(k+1)}$ і подивитися, для яких k послідовність $\mathbb{P}_n(k)$ спадає і для яких k зростає. Довести, що значення $[(n+1)p]$ має найбільшу ймовірність.

Доведемо тепер таке твердження, яке містить характеристичну властивість геометричного розподілу.

Твердження (відсутність післядії)

Для випадкової величини ξ з геометричним розподілом

$$(\forall n \geq 0,\ m \geq 0)\colon\ \mathbb{P}(\{\omega\colon \xi(\omega) = n + m\} | \{\omega\colon \xi(\omega) \geq n\}) = \\ = \mathbb{P}\{\omega\colon \xi(\omega) = m\}.$$

Доведення. За формулою умовної ймовірності

$$\mathbb{P}(\{\omega\colon \xi(\omega) = n + m\} | \{\omega\colon \xi(\omega) \geq n\}) = \\ = \frac{\mathbb{P}(\{\omega\colon \xi(\omega) = n + m\} \cap \{\omega\colon \xi(\omega) \geq n\})}{\mathbb{P}\{\omega\colon \xi(\omega) \geq n\}} = \\ = \frac{\mathbb{P}\{\omega\colon \xi(\omega) = n + m\}}{\mathbb{P}\{\omega\colon \xi(\omega) \geq n\}} = \frac{pq^{n+m}}{\sum_{k=n}^{+\infty} pq^k} = pq^m = \mathbb{P}\{\omega\colon \xi(\omega) = m\}.$$

Якщо для випадкової величини, що набуває лише цілочисельні невід'ємні значення позначити $p_k = \mathbb{P}\{\omega\colon \xi(\omega) = k\}$, то в загальному властивість відсутності післядії з останнього твердження можна описати за допомогою рівності

$$p_m = \frac{p_{n+m}}{\sum\limits_{k=n}^{+\infty} p_k} \quad (n, m \geq 0) \text{ — властивість відсутності післядії.}$$

Нескладно тепер переконатись, що єдиний розподіл, який визначає ця рівність — геометричний.

Твердження

Якщо розподіл $\mathbb{P}\{\xi = k\} = p_k \ (k \geq 0)$ цілочисельної невід'ємної випадкової величини задовольняє умову відсутності післядії, то розподіл геометричний.

Доведення. Справді, виберемо $n = 1$. Тоді, $p_{m+1} = p_m(1 - p_0)$, оскільки $\sum_{k=1}^{+\infty} p_k = 1 - p_0$. Ввівши позначення $q = 1 - p_0$, звідси отримаємо $p_k = p_0 \cdot q^k \ (k \geq 0)$.

2.10. Граничні теореми у схемі Бернуллі: закон подій, що трапляються рідко

Теорема (Пуассона)

В рамках схеми Бернуллі $p = p(n)$, якщо послідовність (a_n), $a_n = np$, — обмежена, то для кожного $k \geq 0$

$$\lim_{n\to+\infty} \frac{P_n(k)}{\frac{a_n^k}{k!}e^{-a_n}} = 1,$$

де $P_n(k) = \mathbb{P}\{\omega\colon m_n(\omega) = k\}$ — ймовірність того, що при n випробуваннях кількість успіхів m_n дорівнює k.

З теореми Пуассона нескладно отримуємо такий наслідок.

Наслідок

Якщо $\lim\limits_{n\to\infty} a_n = a$, то $(\forall k \geq 0)\colon \quad \lim\limits_{n\to+\infty} P_n(k) = \dfrac{a^k}{k!}e^{-a}$.

В наслідку фігурує розподіл Пуассона. Власне, скажемо, що цілочисельна невід'ємна випадкова величина ξ має розподіл Пуассона з параметром $\lambda > 0$ ($\xi \in \mathbb{P}\mathrm{oiss}(\lambda)$), якщо

$$\mathbb{P}\{\omega\colon \xi(\omega) = k\} = \frac{\lambda^k}{k!}e^{-\lambda} \quad (k \geq 0).$$

Доведення теореми Пуассона. Пригадаємо, що

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

Застосовуючи для $n!$ і $(n-k)!$ формулу Стірлінга $s! = \sqrt{2\pi s}\left(\frac{s}{e}\right)^s e^{\frac{\theta_s}{12s}}$, де $0 < |\theta_s| < 1$, отримаємо

$$P_n(k) = \frac{(a_n)^k e^{-a_n}}{k!} \cdot u_{n,k}, \text{ де } u_{n,k} = \frac{e^{np}q^{n-k}\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{\theta_n}{12n} - \frac{\theta_{n-k}}{12(n-k)}}}{n^k\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k}}.$$

Все буде доведено, якщо доведемо, що $\lim\limits_{n\to+\infty} u_{n,k} = 1$.

Але $a_n = np$ — обмежена величина, тому $p \to 0$ $(n \to +\infty)$ і, отже,

$$q = 1 - p \to 1, \quad np^2 \to 0 \quad (n \to +\infty).$$

При $n \to +\infty$ маємо $n - k \to +\infty$, тому для фіксованого $k \geq 0$

$$u_{n,k} \sim \frac{e^{np} q^{n-k} e^{-k}}{(1 - \frac{k}{n})^n} \left(1 - \frac{k}{n}\right)^k \sim e^{np} q^n \quad (n \to +\infty).$$

Покажемо, що $np + n \ln q \to 0$ $(n \to +\infty)$. Запишемо для $y = \ln(1 - t)$ формулу Тейлора з залишковим членом у формі Пеано: $\ln(1 - t) = -t + O(t^2)$ $(t \to 0)$. Тоді

$$np + n \ln q = np + n\left(-p + O(p^2)\right) = O(n \cdot p^2) = o(1) \quad (n \to +\infty).$$

В рамках наслідку з теореми Пуассона Ю.В. Прохоров довів, що для всіх $k \geq 0$ при $n \to +\infty$

$$\sum_{k=0}^{n} |P_n(k) - \frac{a^k}{k!} e^{-a}| \leq \frac{2a}{n} \min\{2, a\}.$$

2.11. Граничні теореми у схемі Бернуллі: локальна Лапласа, інтегральна Муавра-Лапласа. Нормальний (гаусів) розподіл

Теорема (локальна Лапласа)

Якщо в рамках схеми Бернуллі $b_n = npq \to +\infty$, то для кожної фіксованої сталої $c \in (0, +\infty)$ при $(n \to +\infty)$ для всіх $k \in \mathbb{Z}_+$ таких, що $|x_{n,k}| \leq c$, де $x_{n_k} = \frac{k - a_n}{\sqrt{b_n}}$, $a_n = np$,

$$P_n(k) = (1 + o(1)) \frac{1}{\sqrt{2\pi b_n}} e^{-\frac{1}{2} x_{n,k}^2}.$$

Доведення. Позначимо $\alpha = \alpha_{n,k} = \frac{k}{n}$, $\beta = 1 - \alpha$. Тоді, для всіх $n \geq 0, 0 \leq k \leq n$, як і вище, за формулою Стірлінга отримуємо

$$P_n(k) = \frac{1}{\sqrt{2\pi\alpha\beta n}} \left(\frac{p}{\alpha}\right)^{\alpha n} \left(\frac{q}{\beta}\right)^{\beta n} \cdot e^{\theta} = \frac{1}{\sqrt{2\pi\alpha\beta n}} \cdot e^{\theta + n \cdot h(\alpha)},$$

де $\theta = \frac{1}{12 \cdot n}(\theta_n - \frac{\theta_{\alpha n}}{\alpha} - \frac{\theta_{\beta n}}{\beta})$, $h(\alpha) = \alpha(\ln p - \ln \alpha) + \beta(\ln q - \ln \beta)$. Розглянемо тепер функцію
$h(t) = t(\ln p - \ln t) + (1 - t)(\ln q - \ln(1 - t))$ як функцію від $t \in (0, 1)$. Зауважимо, що $h(p) = h^{'}(p) = 0$,
$h^{''}(t) = -\dfrac{1}{t(1 - t)} < 0$ $(t \in (0, 1))$. Записуючи для функції $h(t)$ формулу Тейлора в околі точки $t = p$ з залишковим членом у формі Лагранжа, отримаємо

$$h(\alpha) = \frac{(\alpha - p)^2}{2} h^{''}(\alpha^*) = -\frac{(\alpha - p)^2}{2\alpha^*(1 - \alpha^*)},$$

де α^* — деяка проміжна точка між p і α.

Зауважимо тепер, що $\alpha = p + x_{n,k}\sqrt{\frac{pq}{n}}$, тому для всіх n, k таких що $|x_{n,k}| \le c$, з умови $b_n \to +\infty$ отримуємо, що $\alpha = p + px_{n,k}\frac{q}{\sqrt{b_n}} = p + o(1)$, $\beta = q + qx_{n,k}\frac{p}{\sqrt{b_n}} = q + o(1)$ $(n \to +\infty)$, а, отже, $\alpha^*(1-\alpha^*) = pq + o(1)$ $(n \to +\infty)$. Звідси, $\alpha\beta n \sim b_n$,

$$nh(\alpha) = -n \cdot \frac{(\alpha - p)^2}{2\alpha^*(1-\alpha^*)} = -\frac{x_{n,k}^2}{2} + o(1) \quad (n \to +\infty).$$

Оскільки для всіх k, що задовольняють умову $|x_{n,k}| \le c$, маємо

$$k \ge a_n - c\sqrt{b_n} \ge b_n - c\sqrt{b_n} \to +\infty \ (n \to +\infty),$$

а також

$$n - k \ge nq - c\sqrt{b_n} \ge b_n - c\sqrt{b_n} \to +\infty \ (n \to +\infty),$$

то $\theta = o(1)$ $(n \to +\infty)$ і, отже, остаточно отримаємо потрібне асимптотичне співвідношення.

$$P_n(k) = (1 + o(1))\frac{1}{\sqrt{2\pi b_n}} e^{-\frac{x_{n,k}^2}{2}} \quad (n \to +\infty).$$

Зауважимо, що для функції h з доведення локальної теореми Лапласа $h(p) = h'(p) = 0$, $h''(t) = -\dfrac{1}{t(1-t)} < 0$ $(t \in (0,1))$.

Тому, точка $t = p$ — єдина точка, у якій h' перетворюється в нуль, а також h' — спадна функція на $(0,1)$. Звідси, $h'(t) > 0$ $(t \in (0,p))$, $h'(t) < 0$ $(t \in (p,1))$. Крім цього, $t = p$ — точка максимуму функції $h(t)$.

Вправа. Провівши попередні міркування дещо акуратніше, довести, що еквівалетність в останньому співвідношення рівномірна за всіма k такими, що $|x_{n,k}| \le c$, для кожної фіксованої сталої $c < +\infty$, тобто

$$\lim_{n\to+\infty} \sup\left\{\left|P_n(k) - \frac{1}{\sqrt{2\pi b_n}} e^{-\frac{1}{2}x_{n,k}^2}\right| : k \in \mathbb{Z}_+,\ |x_{n,k}| \le c\right\} = 0.$$

Наступну теорему отримаємо пізніше у вигляді наслідку з центральної граничної теореми.

Теорема (інтегральна Муавра-Лапласа)

Нехай у рамках схеми Бернуллі m_n — кількість успіхів в n незалежних випробуваннях і $b_n = npq \to +\infty$ $(n \to +\infty)$. Тоді

$$(\forall a, b,\ a < b)\colon \quad \lim_{n\to+\infty} \mathbb{P}\left\{\omega\colon\ a \le \frac{m_n - a_n}{\sqrt{b_n}} < b\right\} = \Phi(b) - \Phi(a),$$

де $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$

Відзначимо, що оскільки

$$x_{n,k+1} - x_{n,k} = \frac{1}{\sqrt{b_n}} \quad \text{і} \quad \mathbb{P}\left\{\omega\colon\ a \le \frac{m_n - a_n}{\sqrt{b_n}} < b\right\} = \sum_{a \le x_{n,k} < b} P_n(k),$$

то за допомогою твердження з щойно сформульованої вправи, вибираючи $c = \max\{|a|, |b|\}$, неважко за означенням інтеграла Рімана отримати доведення цієї теореми.

У інтегральній теоремі Муавра-Лапласа $\Phi(x)$ — стандартний нормальний (або гаусів) розподіл.
Скажемо, що випадкова величина $\xi \in \mathcal{N}(0,1)$, тобто має стандартний нормальний розподіл, якщо $F_\xi(x) \equiv \Phi(x)$.
Скажемо, що випадкова величина η має нормальний розподіл з параметрами $a \neq 0,\ \sigma^2 > 0$ ($\eta \in \mathcal{N}(a, \sigma^2)$), якщо випадкова величина $\xi = \dfrac{\eta - a}{\sigma}$ має гаусів розподіл.

Вправи. Переконатись, що: 1. $\Phi(x)$, справді, функція розподілу.
2. Якщо $\eta \in \mathcal{N}(a, \sigma^2)$, то, відповідно, її функція розподілу і щільність

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

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

3. Якщо $\xi \in \mathcal{N}(0,1)$, то $-\xi \in \mathcal{N}(0,1)$. Довести. Такі випадкові величини називаються симетричними.

2.12. Числові характеристики випадкових величин: математичне сподівання і дисперсія

Нехай $(\Omega, \mathcal{A}, \mathbb{P})$ — ймовірнісний простір, $\xi(\omega)$ — випадкова величина.

Математичним сподіванням випадкової величини $\xi(\omega)$ називають

$$M\xi \stackrel{def}{=} \int_{\Omega} \xi(\omega)\mathbb{P}(d\omega),$$

дисперсією випадкової величини $\xi(\omega)$ називають

$$D\xi \stackrel{def}{=} M(\xi - M\xi)^2,$$

стандартом або середньоквадратичним відхиленням називають

$$\sigma \stackrel{def}{=} \sqrt{D\xi}.$$

Зауважимо, що з властивостей інтеграла випливає, що якщо ξ — невід'ємна випадкова величина, то її математичне сподівання завжди існує, при цьому $M\xi \in [0, +\infty]$. Зокрема, якщо у довільної випадкової величини існує скінченне математичне сподівання, то у неї існує дисперсія $D\xi \in [0, +\infty]$, позаяк $(\xi - M\xi)^2$ — невід'ємна випадкова величина.
Якщо розподіл ξ має щільність f_ξ, то

$$M\xi = \int_{-\infty}^{+\infty} x f_\xi(x) dx, \quad D\xi = \int_{-\infty}^{+\infty} (x - M\xi)^2 f_\xi(x) dx.$$

У випадку, коли ξ — дискретна випадкова величина з розподілом $p_n = \mathbb{P}\{\omega \colon \xi(\omega) = x_n\}$, то

$$M\xi = \sum\nolimits_n x_n p_n, \quad D\xi = \sum\nolimits_n (x_n - M\xi)^2 p_n.$$

2.13. Властивості математичного сподівання

Наступні три властивості математичного сподівання, насправді, є властивостями відповідного інтеграла

1. Якщо випадкова величина ξ є майже напевно (м.н.) стала, тобто, $\mathbb{P}\{\omega : \xi(\omega) = a\} = 1$, то $M\xi = a$.
Нехай $B = \{\omega : \xi(\omega) \neq a\}$. Тоді $\mathbb{P}(B) = 0$ і, отже,

$$M\xi = \int_{\Omega\setminus B} \xi(\omega)\mathbb{P}(d\omega) + \int_B \xi(\omega)\mathbb{P}(d\omega) = a \cdot \mathbb{P}(\overline{B}) = a.$$

Означення

Скрізь надалі, якщо ймовірність події A дорівнює 1, то ми кажемо, що подія A відбувається майже напевно (м.н.).

2. (адитивність). Нехай $(\xi_n)_{n=1}^N$, $N < +\infty$ — послідовність випадкових величин. Тоді

$$M\left(\sum\nolimits_{n=1}^N \xi_n\right) = \sum\nolimits_{n=1}^N M\xi_n.$$

3. (лінійність). Нехай ξ — випадкова величина. Тоді $(\forall \alpha \in \mathbb{R})$: $M(\alpha\xi) = \alpha M\xi$.

4. (монотонність). Для будь-яких двох випадкових величин ξ, η з того, що $\xi \geq \eta$ (м.н.) (тобто, $\mathbb{P}\{\omega\colon \xi(\omega) \geq \eta(\omega)\} = 1$) випливає, що $M\xi \geq M\eta$.
Справді, оскільки $\xi - \eta \geq 0$ (м.н.), то залишається скористатись невід'ємністю інтеграла від невід'ємної випадкової величини і лінійністю.

5. Якщо випадкова величина ξ така, що $M\xi \neq \infty$, то $|M\xi| \leq M|\xi|$.
Справді, $-|\xi| \leq \xi \leq |\xi|$. тому за попередньою властивістю отримуємо потрібну нерівність.

6. Якщо $A \in \mathcal{A}$, то $M\chi_A = \mathbb{P}(A)$, де $\chi_A(\omega) = \begin{cases} 1, & \text{якщо } \omega \in A, \\ 0, & \text{якщо } \omega \notin A \end{cases}$
— характеристична функція події A.

7. (Нерівність Маркова). Нехай ξ — невід'ємна випадкова величина. Тоді

$$(\forall a > 0)\colon\ \mathbb{P}\{\omega : \xi(\omega) \geq a\} \leq \frac{M\xi}{a}$$

Доведення. Позначимо $A = \{\omega\colon \xi(\omega) \geq a\}$. Тоді,

$$\mathbb{P}(A) = \int_A \mathbb{P}(d\omega) \leq \int_A \frac{\xi(\omega)}{a}\mathbb{P}(d\omega) \leq \frac{1}{a}\int_\Omega \xi(\omega)\mathbb{P}(d\omega) = \frac{1}{a}M\xi.$$

2.14. Властивості дисперсії

1. Для кожної випадкової величини зі скінченним математичним сподіванням $D\xi = M\xi^2 - (M\xi)^2$.

Доведення.

Використовуємо адитивність і лінійність математичного сподівання

$$D\xi = M(\xi^2 - 2\xi M\xi + (M\xi)^2) = M\xi^2 + M(-2\xi M\xi) + M(M\xi)^2 =$$
$$= M\xi^2 - 2M\xi M\xi + (M\xi)^2 = M\xi^2 - (M\xi)^2.$$

□

Звідси, в часткових випадках дискретної і абсолютно неперервної випадкової величин отримуємо такі формули для обчислення їхньої дисперсії: ξ — дискретна випадкова величина, тоді

$$D\xi = \sum\nolimits_n x_n^2 p_n - \left(\sum\nolimits_n x_n p_n\right)^2, \quad p_n = \mathbb{P}\{\omega\colon \xi(\omega) = x_n\}, \ \sum\nolimits_n p_n = 1,$$

ξ — абсолютно неперервна з щільністю розподілу f_ξ, тоді

$$D\xi = \int_{-\infty}^{+\infty} x^2 f_\xi(x)dx - \left(\int_{-\infty}^{+\infty} x f_\xi(x)dx\right)^2.$$

2. Для довільної випадкової величини зі скінченною дисперсією

$$|M\xi| \le \sqrt{M\xi^2}.$$

Справді, $0 \le D\xi = M\xi^2 - (M\xi)^2$. Звідси отримуємо потрібну нерівність.

3. Випадкова величина ξ є м.н. сталою (тобто, $\xi(\omega) = c$ м.н.) $\Longleftrightarrow D\xi = 0$.

Доведення.

Послідовно маємо $\xi^2(\omega) = c^2$ (м.н.) $\Longrightarrow$ $M\xi^2 = c^2$. Тому $D\xi = c^2 - (c)^2 = 0$.
Навпаки, $D\xi = M(\xi - M\xi)^2 = 0$. Тоді $(\xi - M\xi)^2 = 0$ м.н.
Остаточно, $\xi = M\xi$ м.н. $\square$

З означення дисперсії і властивості лінійності математичного сподівання замість лінійності, для дисперсії отримуємо властивість однорідності порядку 2.

4. $(\forall \alpha \in \mathbb{R})$: $D(\alpha\xi) = \alpha^2 D\xi$.

На відміну від математичного сподівання, для того, щоб дисперсія стала адитивною, від випадкових величин потрібно вимагати деяких властивостей, які вони повинні задовольняти "колективно". Зокрема, виконується таке твердження.

5. $D(\xi + \eta) = D\xi + D\eta \iff M(\xi\eta) = M\xi M\eta$.

Доведення.

Розписуючи детально $D(\xi + \eta)$, отримуємо

$$D(\xi+\eta) = M(\xi+\eta - M(\xi+\eta))^2 = M((\xi - M\xi) + (\eta - M\eta))^2 =$$
$$= M(\xi - M\xi)^2 + M(\eta - M\eta)^2 + 2M((\xi - M\xi)(\eta - M\eta)) =$$
$$= D\xi + D\eta + 2(M(\xi\eta) - M\xi M\eta).$$

□

Звідси, зокрема, безпосередньо випливає таке твердження.

6. Якщо $\eta = c$ м.н., то $D(\xi + \eta) = D\xi$ для кожної випадкової величини ξ.

Справді, $M(\eta\xi) = \int_\Omega \eta\xi \mathbb{P}(d\omega) = M(c\xi) = cM\xi = M\eta M\xi$, тому за властивістю 5. $D(\xi + \eta) = D\xi + D\eta$, але за властивістю 3. $D\eta = 0$.

Означення (коваріації і некорельованих випадкових величин)

Коваріацією випадкових величин ξ, η називають величину

$$\mathrm{cov}(\xi, \eta) \stackrel{def}{=} M((\xi - M\xi)(\eta - M\eta)).$$

Випадкові величини ξ, η називаються некорельованими, якщо

$$\mathrm{cov}(\xi, \eta) = 0.$$

Якщо ж $\mathrm{cov}(\xi, \eta) \neq 0$, то випадкові величини ξ, η називаються корельованими.

Зауважимо, що поняття коваріації та корельованості (некорельованості) має зміст лише для випадкових величин ξ, η інтегровних з квадратом на Ω, тобто, лише у випадку, коли існують скінченні дисперсії $D\xi, D\eta$.
З доведення вправи 5. отримуємо такі твердження.

7. $D(\xi+\eta) = D\xi + D\eta \iff$ випадкові величини ξ, η некорельовані, тобто $\operatorname{cov}(\xi, \eta) = 0$.

8. Випадкові величини ξ, η некорельовані $\iff M(\xi\eta) = M\xi M\eta$.

9. Якщо (ξ_j) — послідовність попарно некорельованих випадкових величин, тобто, $\operatorname{cov}(\xi_j, \xi_k) = 0$ $(j \neq k)$, то $D\left(\sum_{k=1}^{n} \xi_k\right) = \sum_{k=1}^{n} D\xi_k$.

Доведення.

Доведення проведемо з використанням методу математичної індукції. Базою індукції є властивість 7. Для того, щоб можна було провести крок індукції, залишається перевірити некорельованість випадкових величин ξ_n та $\eta = \sum_{j=1}^{n-1} \xi_j$ у випадку, коли $(\xi_j)_{j=1}^n$ — попарно некорельовані, за властивістю 8. досить перевірити, що $M(\xi_n\eta) = M\xi_n M\eta$. Остання ж рівність виконується. Справді,

$$M(\xi_n\eta) = M\Big(\sum_{j=1}^{n-1}(\xi_j\xi_n)\Big) =$$

$$= \sum_{j=1}^{n-1} M(\xi_j\xi_n) = M\xi_n\Big(\sum_{j=1}^{n-1} M\xi_j\Big) = M\xi_n M\eta.$$

Оскільки ξ_n та $\eta = \sum_{j=1}^{n-1} \xi_j$ — некорельовані, то, скориставшись властивістю 7. і припущенням індукції, остаточно отримуємо

$$D\Big(\sum_{j=1}^{n} \xi_j\Big) = D(\xi_n + \eta) = D\xi_n + D\eta = \sum_{j=1}^{n} D\xi_j.$$

$\square$

2.15. Приклади обчислення дисперсії деяких розподілів

1. Схема Бернуллі: геометричний розподіл. Нехай випадкова величина ξ має геометричний розподіл, тобто $p_n = \mathbb{P}\{\omega\colon \xi(\omega) = n\} = pq^n$ $(n \in \mathbb{Z}_+)$. Тоді

$$M\xi = \sum\nolimits_{n=0}^{+\infty} np_n = \sum\nolimits_{n=0}^{+\infty} nq^n p = q\sum\nolimits_{n=1}^{+\infty} nq^{n-1}p =$$

$$= q\Big(\sum\nolimits_{n=0}^{+\infty} q^n\Big)' p = qp\Big(\frac{1}{1-q}\Big)' = \frac{qp}{(1-q)^2} = \frac{qp}{p^2} = \frac{q}{p}.$$

Подібно

$$M\xi^2 = \sum\nolimits_{n=0}^{+\infty} n^2 q^n p = pq\Big(q\Big(\frac{1}{1-q}\Big)'\Big)' = \frac{(1-p)(2-p)}{p^2}.$$

Тому $D\xi = qp^{-2}$.

2. Схема Бернуллі: біномний розподіл. Оскільки випадкова величина m_n — кількість У в n незалежних випробуваннях записується у вигляді $m_n = \sum_{j=1}^n \xi_{j,n}$, то зауважуючи, що $M\xi_{j,n} = 1 \cdot p + 0 \cdot q = p$, отримаємо

$$Mm_n = \sum_{j=1}^n M\xi_{j,n} = np = a_n.$$

Якщо доведемо, що $(\xi_{j,n})_{j=1}^n$ — попарно некорельовані, то, обчислюючи дисперсію
$D\eta_{j,n} = M\eta_{j,n}^2 - (M\eta_{j,n})^2 = 1 \cdot \mathbb{P}\{\eta_{j,n}^2 = 1\} + 0 \cdot \mathbb{P}\{\eta_{j,n}^2 = 0\} - p^2 =$
$= 1 \cdot p - p^2 = pq$, отримаємо

$$Dm_n = \sum_{j=1}^n D\xi_{j,n} = npq = b_n.$$

Для, того, щоб встановити некорельованість $\xi_{k,n}, \xi_{j,n}$ $(k \neq j)$ досить перевірити, що $M\eta = M\xi_{k,n}M\xi_{j,n}$, де $\eta = \xi_{k,n}\xi_{j,n}$. Зауважимо, що $\mathbb{P}\{\omega\colon \eta(\omega) = 1\} = \mathbb{P}(\text{У})\mathbb{P}(\text{У}) = p^2$ (тут ми скористались незалежністю випробовувань, оскільки з неї випливає незалежність відповідних випадкових величин $\xi_{k,n}, \xi_{j,n}$) та $\mathbb{P}\{\omega\colon \eta(\omega) = 0\} = 1 - p^2$. Тому, $M\eta = p^2$. Але, вище ми обчислили, що $M\xi_{k,n} = M\xi_{j,n} = p$.

Вправи. 1. Переконатись, що випадкові величини $\xi_{k,n}, \xi_{j,n}$ $(k \neq j)$ — незалежні. Вище було доведено, що звідси випливає некорельованість цих випадкових величин.
2. Переконатись, що правильною є загальніша імплікація: якщо дискретні випадкові величини ξ_1, ξ_2 — незалежні, то вони й некорельовані (насправді, те, що це твердження правильне в загальному — для довільної пари незалежних випадкових величин, ми доведемо далі).

3. Розподіл Пуассона. Нехай $\xi \in \mathcal{P}oiss(\lambda)$, тобто, має розподіл Пуассона з параметром $\lambda > 0$: $\mathbb{P}\{\omega\colon \xi(\omega) = n\} = \frac{\lambda^n}{n!}e^{-\lambda}$, $n \in \mathbb{Z}_+$. Тоді

$$M\xi = \sum\nolimits_{n=0}^{\infty} n\frac{\lambda^n}{n!}e^{-\lambda} = \lambda\sum\nolimits_{n=1}^{+\infty} n\frac{\lambda^{n-1}}{n!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!} = \lambda.$$

Обчислимо дисперсію $D\xi$ в розподілі Пуассона. Оскільки

$$M\xi^2 = \sum\nolimits_{n=0}^{+\infty} n^2\frac{\lambda^n}{n!}e^{-\lambda} = \lambda\sum\nolimits_{n=1}^{+\infty} n^2\frac{\lambda^{n-1}}{n!}e^{-\lambda} =$$

$$\lambda e^{-\lambda}\left(\sum\nolimits_{n=1}^{+\infty}\frac{\lambda^n}{(n-1)!}\right)' =$$

$$= \lambda e^{-\lambda}\left(\lambda e^{\lambda}\right)' = \lambda(\lambda+1).$$

Тому $D\xi = M(\xi^2) - (M\xi)^2 = \lambda(\lambda+1) - \lambda^2 = \lambda.$

4. Рівномірний розподіл. Нехай випадкова величина ξ має рівномірний розподіл на проміжку $[a, b]$. Тоді

$$M\xi = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2}.$$

А, оскільки,

$$M\xi^2 = \int_{-\infty}^{+\infty} x^2 f_\xi(x) dx = \frac{1}{b-a} \int_a^b x^2 dx = \frac{a^2+b^2+ab}{3},$$

то $D\xi = M(\xi)^2 - (M\xi)^2 = \dfrac{(b-a)^2}{12}$.

5. Нормальний розподіл. Обчислимо спочатку математичне сподівання і дисперсію для стандартного нормального розподілу $\mathcal{N}(0,1)$. Нехай $\xi \in \mathcal{N}(0,1)$. Тоді,

$$M\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}} dx = 0,$$

позаяк підінтегральна функція непарна. Тому, для обчислення дисперсії маємо

$$D\xi = M(\xi)^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx =;$$

$$-\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1.$$

Для випадкової величини $\eta \in \mathcal{N}(a, \sigma^2)$, оскільки $\eta = a + \xi\sigma$, де $\xi \in \mathcal{N}(0,1)$, з властивостей математичного сподівання і дисперсії отримаємо

$$M\eta = a + \sigma M\xi = a, \quad D\eta = \sigma^2 D\xi = \sigma^2.$$

6. Розподіл Коші. Чи існує математичне сподівання $M\xi$ для випадкової величини $\xi \in \mathcal{K}(0,1)$, тобто, з розподілом Коші? (Переконайтесь, що у звичайному сенсі воно не існує. Чи можна $M\xi$ визначити в сенсі головного значення інтеграла?)

Твердження

Нехай $\chi_1, \chi_2, \ldots, \chi_n$ — індикатори незалежних подій $A_1, A_2, \ldots, A_n$, відповідно. Тоді

$$M(\chi_1\chi_2\cdots\chi_n) = M\chi_1 M\chi_2 \cdots M\chi_n.$$

Доведення.

Розглянемо індикатор події $A \stackrel{def}{=} \bigcap\limits_{j=1}^{n} A_j$. Тоді $\chi_A = \chi_1\chi_2\cdots\chi_n$ і, отже,

$$M(\chi_1\chi_2\ldots\chi_n) = M\chi_A = \mathbb{P}(A) = \Big[\text{з незалежності подій } (A_j)\Big] =$$

$$= \prod_{j=1}^{n} \mathbb{P}(A_j) = \prod_{j=1}^{n} \int_{A_j} \chi_j \mathbb{P}(d\omega) = \prod_{j=1}^{n} \int_{\Omega} \chi_j \mathbb{P}(d\omega) = \prod_{j=1}^{n} M\chi_j.$$

$\square$

Означення (простої випадкової величини)

Випадкову величину, множина значень якої — скінченна, називатимемо простою випадковою величиною.

Вправа. Довести, що випадкова величина ξ — проста, тоді і тільки тоді, коли її можна подати у вигляді

$$\xi = \sum\nolimits_{s=1}^{m} a_s \chi_s, \quad 1 \le m < +\infty, \ a_s \in \mathbb{R},$$

де χ_s — індикатор події $A_s = \{\omega \colon \xi(\omega) = a_s\}$.

Твердження

Нехай $\xi_1, \xi_2, \ldots, \xi_n$ — незалежні прості випадкові величини. Тоді

$$M(\xi_1 \xi_2 \cdots \xi_n) = M\xi_1 M\xi_2 \cdots M\xi_n,$$

зокрема, (ξ_j) — послідовність попарно некорельованих випадкових величин.

Доведення. Нехай $\xi_j = \sum_{s=1}^{m_j} a_{s,j} \chi_{s,j}$, де $\{a_{s,j} \colon 1 \le s \le m_j\}$ — множина значень ξ_j, $\chi_{s,j}$ — індикатор події $A_{s,j} = \{\omega \colon \xi_j(\omega) = a_{s,j}\}$. Тоді

$$M(\xi_1 \xi_2 \ldots \xi_n) = M\Big(\prod_{j=1}^{n} \Big(\sum\nolimits_{s_j=1}^{m_j} a_{s_j,j} \chi_{s_j,j}\Big)\Big) =$$

$$= \sum\nolimits_{s_1=1,\ldots,s_n=1}^{m_1,\ldots,m_n} M\Big(\prod_{j=1}^{n} a_{s_j,j} \chi_{s_j,j}\Big) = \sum\nolimits_{s_1=1,\ldots,s_n=1}^{m_1,\ldots,m_n} \prod_{j=1}^{n} a_{s_j,j} M\Big(\prod_{j=1}^{n} \chi_{s_j,j}\Big).$$

Але незалежність послідовності випадкових величин (ξ_j) еквівалентна до незалежності послідовності σ-алгебр $(\mathcal{A}_{\xi_j})$, породжених цими випадковими величина. Враховуючи тепер, що $(\forall j,\ 1 \le j \le n)(\forall s,\ 1 \le s \le m_j)\colon A_{s,j} \in \mathcal{A}_{\xi_j}$, звідси негайно отримаємо, що індикатори $(\chi_{s_j,j})_{j=1}^n$ — незалежні випадкові величини для будь-яких наборів $(s_1, \ldots, s_n)$, $1 \le s_j \le m_j$ $(1 \le j \le n)$. Тому, за твердженням математичне сподівання індикаторів незалежних подій отримуємо, що $M\Big(\prod\limits_{j=1}^n \chi_{s_j,j}\Big) = \prod\limits_{j=1}^n M\chi_{s_j,j}$. Отже, остаточно,

$$M(\xi_1\xi_2\ldots\xi_n) = \sum\nolimits_{s_1=1}^{m_1} \cdots \sum\nolimits_{s_n=1}^{m_n} \Big(\prod_{j=1}^n a_{s_j,j} \prod_{j=1}^n M\Big(\chi_{s_j,j}\Big)\Big) =$$

$$= \sum\nolimits_{s_1=1,\ldots,s_n=1}^{m_1,\ldots,m_n} \Big(\prod_{j=1}^n a_{s_j,j} M\chi_{s_j,j}\Big) = \prod_{j=1}^n \Big(\sum\nolimits_{s_j=1}^{m_j} a_{s_j,j} M\chi_{s_j,j}\Big) = \prod_{j=1}^n M\xi_j.$$

2.16. Нерівність Чебишова і Закон великих чисел (ЗВЧ)

Почнемо з наступного важливого доведеного П.Л. Чебишовим твердження, яке, як це часто у таких випадках буває, має нескладне доведення.

Твердження (Нерівність Чебишова)

Нехай ξ — випадкова величина, що має скінченне математичне сподівання $M\xi \neq \infty$. Тоді,

$$(\forall \varepsilon > 0)\colon\ \mathbb{P}\{\omega : |\xi(\omega) - M\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}.$$

Доведення.

Застостосовуючи нерівність Маркова до невід'ємної випадкової величини $\eta(\omega) = (\xi(\omega) - M\xi)^2$, оскільки $M\eta = D\xi$, отримаємо

$$\mathbb{P}\{\omega : |\xi(\omega) - M\xi| \geq \varepsilon\} = \mathbb{P}\{\omega : (\xi(\omega) - M\xi)^2 \geq \varepsilon^2\} \leq \frac{M\eta}{\varepsilon^2} = \frac{D\xi}{\varepsilon^2}.$$

$\square$

Твердження (узагальнена нерівність Чебишова)

Нехай $X \in \mathcal{B}(\mathbb{R})$ — множина значень випадкової величини ξ, а функція $h\colon X \to \mathbb{R}_+$ монотонно неспадна. Тоді

$$(\forall a \in X)\colon\ \mathbb{P}\{\omega\colon \xi(\omega) \geq a\} \leq \frac{Mh(\xi)}{h(a)}$$ — загальна нерівність Чебишова

Доведення.

З монотонності $h(x) \nearrow$ отримуємо

$$\{\omega : \xi(\omega) \geq a\} \subset \{\omega : h(\xi(\omega)) \geq h(a)\}.$$

Залишається застосувати нерівність Маркова. $\square$

Тепер у нас вже є увесь набір допоміжних фактів для того, щоб довести таку граничну теорему — закон великих чисел (ЗВЧ).

Теорема (Чебишова)

Нехай (ξ_n) — послідовність попарно некорельованих випадкових величин таких, що

$$(\exists C \in (0, +\infty))(\forall n \geq 1)\colon\ D\xi_n \leq C.$$

Тоді

$$(\forall \varepsilon > 0)\colon\quad \lim_{n\to+\infty} \mathbb{P}\left\{\omega\colon \left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k\right| \geq \varepsilon\right\} = 0.$$

Доведення.

Оскільки (ξ_n) — послідовність попарно некорельованих випадкових величин, то $D(\sum_{k=1}^{n}\xi_k) = \sum_{k=1}^{n} D\xi_k$. Тому, застосовуючи нерівність Чебишова, отримуємо

$$0 \leq \frac{1}{n^2} D\left(\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} D\xi_k \leq \frac{1}{n^2}\sum_{k=1}^{n} C = \frac{C}{n} \to 0 \quad (n \to \infty).$$

□

Практично дослівне повторення попередніх міркувань дає змогу отримати таке твердження.

Теорема

Нехай (s_n) — довільна послідовність натуральних чисел така, що $\lim_{n\to+\infty} s_n = +\infty$, а $(\xi_{j,n})_{j=1}^{s_n}$ — послідовність випадкових величин таких, що $\xi_{j,n}$ і $\xi_{k,n}$ некорельовані для кожних $j \neq k$ та будь-якого n. Якщо виконується умова

$$(\exists C \in (0,+\infty))(\forall n \geq 1)(\forall j,\ 1 \leq j \leq s_n)\colon\ D\xi_{j,n} \leq C,$$

то

$$(\forall \varepsilon > 0)\colon\quad \lim_{n\to+\infty} \mathbb{P}\left\{\omega\colon \left|\frac{1}{s_n}\sum_{k=1}^{s_n}\xi_{k,n} - \frac{1}{s_n}\sum_{k=1}^{s_n} M\xi_{k,n}\right| \geq \varepsilon\right\} = 0.$$

Вправа. Довести цю теорему.

Якщо у цій теоремі перейти до протилежних подій, то отримаємо такий наслідок.

Наслідок

Нехай $(\xi_{j,n})_{j=1}^{n}$ — послідовність випадкових величин таких, що $\xi_{j,n}$ і $\xi_{k,n}$ некорельовані для кожних $j \neq k$ та будь-якого n. Якщо виконується умова

Безпосередньо з цього наслідку отримуємо таку теорему Бернуллі.

Наслідок (теорема Бернуллі)

В рамках схеми Бернуллі виконується

$$(\forall \varepsilon > 0)\colon \lim_{n\to+\infty} \mathbb{P}\{\omega\colon |\nu_n - p| < \varepsilon\} = 1,$$

де $\nu_n = \frac{m_n}{n}$ — частота успіхів (частота появи події A) і $p = \mathbb{P}(\text{У})$ в n послідовних незалежних випробуваннях.

Доведення.

Досить пригадати, що $m_n = \sum_{k=1}^{n} \xi_{k,n}$, де $\xi_{k,n}$ — індикатор появи У в k-тому випробуванні у серії з n послідовних випробувань, та, що послідовність $(\xi_{k,n})$ задовольняє умови наслідку 2. При цьому $M\xi_{k,n} = p$ $(1 \le k \le n)$. $\square$

Введемо тепер поняття збіжності послідовності випадкових величин за ймовірністю, як тотожне з поняттям збіжності за мірою $\mathbb{P}$.

Означення (збіжності за ймовірністю)

Послідовність випадкових величин (η_n) називається збіжною за ймовірністю до випадкової величини η, $\eta_n \xrightarrow{\mathrm{P}} \eta \quad (n \to +\infty)$, якщо

$$(\forall \varepsilon > 0)\colon \quad \lim_{n\to+\infty} \mathbb{P}\{\omega\colon |\eta_n(\omega) - \eta(\omega)| \geq \varepsilon\} = 0$$

$$\left(\text{рівносильна умова: } (\forall \varepsilon > 0)\colon \quad \lim_{n\to+\infty} \mathbb{P}\{\omega\colon |\eta_n(\omega) - \eta(\omega)| < \varepsilon\} = 1\right).$$

Наслідок (класична теорема Бернуллі)

Якщо в рамках схеми Бернуллі ймовірність успіху $p(n) \equiv p$ $(n \geq 0)$, тобто — стала, то

$$\nu_n \xrightarrow{\mathrm{P}} p \quad (n \to +\infty).$$

Подібно можна переформулювати решту інших, наведених тут ЗВЧ. На цьому зупинятись не будемо, а сформулюємо ще одне твердження типу ЗВЧ, схема доведення якого у нас по-суті вже є.

Теорема

Нехай $(\xi_{j,n})_{j=1}^{n}$ — послідовність випадкових величин таких, що $\xi_{j,n}$ і $\xi_{k,n}$ некорельовані для кожних $j \neq k$ та будь-якого n. Якщо виконується умова $\lim\limits_{n\to+\infty} \frac{1}{n}\sum_{j=1}^{n} D\xi_{j,n} = 0$, то

$$\frac{1}{n}\sum_{k=1}^{n} \xi_{k,n} - \frac{1}{n}\sum_{k=1}^{n} M\xi_{k,n} \xrightarrow{\mathrm{P}} 0 \quad (n \to +\infty).$$

Вправа. Довести цю теорему.

Відзначимо, ще деякі цікаві і корисні факти, які безпосередньо випливають з нерівностей Чебишова і Маркова.

Вправи. 1. (закон трьох сигм). Нехай ξ — випадкова величина і $\sigma^2 = D\xi$. Тоді

$$\mathbb{P}\{\omega\colon |\xi - M\xi| \geq 3\sigma\} \leq \frac{1}{9} = 0,(1).$$

2. Нехай послідовність випадкових величин (ξ_n) збіжна до випадкової величини ξ в середньому, тобто, $M|\xi_n - \xi| \to 0$ $(n \to +\infty)$. Довести, що тоді:
a) $\xi_n \xrightarrow{\mathrm{P}} \xi$ $(n \to +\infty)$; b) якщо $M\xi_n \neq \infty, M\xi \neq \infty$, то $M\xi_n \to M\xi$ $(n \to +\infty)$. Вказівка. a) Застосувати нерівність Маркова до $|\xi_n - \xi|$.

3. Нехай послідовність випадкових величин (ξ_n) збіжна до випадкової величини ξ в середньоквадратичному, тобто, $M|\xi_n - \xi|^2 \to 0$ $(n \to +\infty)$. Довести, що тоді: a) послідовність (ξ_n) збіжна в середньому, а, отже, і $\xi_n \xrightarrow{\mathrm{P}} \xi$ $(n \to +\infty)$; b) якщо $M\xi_n^2 < +\infty, M\xi^2 < +\infty$, то $M\xi_n \to M\xi$, $M(\xi_n\xi) \to M(\xi)^2$, $M(\xi_n)^2 \to M(\xi)^2$, $D\xi_n \to D\xi$, $(n \to +\infty)$.

Вказівки. a) Застосувати нерівність

$$(M\eta)^2 \leq M(\eta)^2 \text{ до } |\xi_n - \xi|.$$

b) $|M(\xi_n\xi) - M(\xi)^2| \leq M|(\xi_n - \xi)\xi| \leq (M|(\xi_n - \xi)^2)^{1/2}(M(\xi)^2)^{1/2}$,
$|M(\xi_n)^2 - M(\xi)^2| = |M(\xi_n - \xi)(\xi_n + \xi)| \leq$
$(M(\xi_n - \xi)^2)^{1/2}(M(\xi_n + \xi)^2)^{1/2}$,
$M(\xi_n + \xi)^2 = M((\xi_n - \xi) + \xi)^2 = M(\xi_n - \xi)^2 + 2M(\xi_n\xi) - M(\xi^2)$.

2.17. Незалежні випадкові величини

Нехай $(\Omega, \mathcal{A}, \mathbb{P})$ — фіксований ймовірнісний простір.
Повернемось до незалежних випадкових величин. Нагадаємо, що послідовність незалежних випадкових величин це така послідовність, що відповідна послідовність σ-алгебр, породжених елементами послідовності випадкових величин, є послідовністю незалежних σ-алгебр.
Доведемо такий критерій незалежності випадкових величин.

Теорема (критерій незалежності)

(ξ_n) — послідовність незалежних випадкових величини $\Longleftrightarrow$ $(\forall n \in \mathbb{N})(\forall x_j \in \mathbb{R},\ 1 \le j \le n)$:

$$\mathbb{P}\Big(\bigcap_{j=1}^{n}\{\omega : \xi_j(\omega) < x_j\}\Big) = \prod_{j=1}^{n}\mathbb{P}\big(\{\omega : \xi_j(\omega) < x_j\}\big).$$

Доведення.Необхідність. (ξ_n) — послідовність незалежних випадкових величин, якщо $\mathcal{A}_{\xi_j} = \xi_j^{-1}(\mathcal{B}(\mathbb{R}))$ — послідовність незалежних σ-алгебр, тобто

$$(\forall n \in \mathbb{N})(\forall A_j \in \mathcal{A}_{\xi_j}):\ \mathbb{P}\Big(\bigcap_{j=1}^{n} A_j\Big) = \prod_{j=1}^{n} \mathbb{P}(A_j).$$

Але, $A_j = \xi_j^{-1}(B_j)$, $B_j \in \mathcal{B}(\mathbb{R})$. Залишається вибрати $B_j = (-\infty, x_j)$ і зауважити, що у цьому випадку $A_j = \xi_j^{-1}(B_j) = \{\omega : \xi_j(\omega) < x_j\}$.

Достатність. Доведемо спочатку таке твердження.

Твердження

Нехай ξ — випадкова величина, подія $A \in \mathcal{A}$ така, що $\mathbb{P}(\xi^{-1}(-\infty, x) \cap A) = \mathbb{P}(\xi^{-1}(-\infty, x))\mathbb{P}(A)$ для кожного $x \in \mathbb{R}$. Через $\mathcal{M}_0$ позначимо сім'ю множин $B \subset \mathbb{R}$, для яких виконується рівність

$$\mathbb{P}(\xi^{-1}(B) \cap A) = \mathbb{P}(\xi^{-1}(B))\mathbb{P}(A).$$

Тоді $\mathcal{M}_0$ — σ-алгебра, а, отже, $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_0$.

Перевіримо спочатку,що якщо $B_1, B_2 \in \mathcal{M}_0$, $B_1 \cap B_2 = \varnothing$, то $B_1 \cup B_2 \in \mathcal{M}_0$. Справді, використовуючи властивості операції взяття прообразу та несумісність $\xi^{-1}(B_1), \xi^{-1}(B_2)$, послідовно отримуємо

$$\mathbb{P}(\xi^{-1}(B_1 \cup B_2) \cap A) = \mathbb{P}\Big((\xi^{-1}(B_1) \cup \xi^{-1}(B_2)) \cap A\Big) =$$

$$\mathbb{P}\Big((\xi^{-1}(B_1) \cap A) \cup (\xi^{-1}(B_2) \cap A)\Big) = \mathbb{P}\Big(\xi^{-1}(B_1) \cap A\Big) + \mathbb{P}\Big(\xi^{-1}(B_2) \cap A\Big)$$

$$= \Big(\mathbb{P}(\xi^{-1}(B_1)) + \mathbb{P}(\xi^{-1}(B_2))\Big)\mathbb{P}(A) = \mathbb{P}\Big(\xi^{-1}(B_1) \cup \xi^{-1}(B_2)\Big)\mathbb{P}(A) =$$

$$= \mathbb{P}\Big(\xi^{-1}(B_1 \cup B_2)\Big)\mathbb{P}(A),$$

а це і означає, що $B_1 \cup B_2 \in \mathcal{M}_0$.

Зауважимо, що за неперервністю ймовірності $\mathbb{R} \in \mathcal{M}_0$, позаяк

$$\mathbb{P}(\xi^{-1}(\mathbb{R}) \cap A) = \mathbb{P}\Big(\xi^{-1}\Big(\bigcup_{n=1}^{+\infty}(-\infty, n)\Big) \cap A\Big) =$$

$$= \mathbb{P}\Big(\bigcup_{n=1}^{+\infty} \xi^{-1}\Big((-\infty, n)\Big) \cap A\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\xi^{-1}\Big((-\infty, n)\Big) \cap A\Big) =$$

$$= \mathbb{P}(A) \lim_{n\to+\infty} \mathbb{P}\Big(\xi^{-1}\big((-\infty, n)\big)\Big) = \mathbb{P}(A)\mathbb{P}\Big(\bigcup_{n=1}^{+\infty} \xi^{-1}\big((-\infty, n)\big)\Big) =$$

$$= \mathbb{P}(\xi^{-1}(\mathbb{R}))\mathbb{P}(A).$$

Нехай тепер $B \in \mathcal{M}_0$. Переконаємось, що $\overline{B} = \mathbb{R} \setminus B \in \mathcal{M}_0$. Справді,

$$\mathbb{P}(\xi^{-1}(\overline{B}) \cap A) = \mathbb{P}(\xi^{-1}(\mathbb{R} \setminus B) \cap A) = \mathbb{P}\Big((\xi^{-1}(\mathbb{R}) \setminus \xi^{-1}(B)) \cap A\Big) =$$

$$= \mathbb{P}\Big((\xi^{-1}(\mathbb{R}) \cap A) \setminus (\xi^{-1}(B) \cap A)\Big) = \mathbb{P}\Big(\xi^{-1}(\mathbb{R}) \cap A\Big) - \mathbb{P}\Big(\xi^{-1}(B) \cap A\Big) =$$

$$= \Big(\mathbb{P}(\xi^{-1}(\mathbb{R})) - \mathbb{P}(\xi^{-1}(B))\Big)\mathbb{P}(A) = \mathbb{P}\Big(\xi^{-1}(\mathbb{R}) \setminus \xi^{-1}(B)\Big)\mathbb{P}(A) =$$

$$= \mathbb{P}(\xi^{-1}(\overline{B}))\mathbb{P}(A).$$

Тепер вже нескладно переконуємось, що $B_1, B_2 \in \mathcal{M}_0 \Longrightarrow B_1 \cup B_2 \in \mathcal{M}_0$. Справді,

$$\mathbb{P}(\xi^{-1}(B_1 \cup B_2) \cap A) = \mathbb{P}\Big((\xi^{-1}(B_1) \cap A) \cup (\xi^{-1}(B_2) \cap A)\Big) =$$

$$= \mathbb{P}\Big(\xi^{-1}(B_1) \cap A\Big) + \mathbb{P}\Big(\xi^{-1}(B_2) \cap A\Big) - \mathbb{P}\Big(\xi^{-1}(B_1) \cap \xi^{-1}(B_2) \cap A\Big).$$

Але,

$$\mathbb{P}\Big(\xi^{-1}(B_1)\cap\xi^{-1}(B_2)\cap A\Big)=\mathbb{P}\Big(\xi^{-1}(B_1\cap B_2)\cap A\Big)=$$
$$\mathbb{P}\Big(\xi^{-1}(B_1\cap B_2)\Big)\mathbb{P}(A),$$

тому остаточно

$$\mathbb{P}(\xi^{-1}(B_1\cup B_2)\cap A)=\Big(\mathbb{P}(\xi^{-1}(B_1))+\mathbb{P}(\xi^{-1}(B_2))-$$
$$-\mathbb{P}(\xi^{-1}(B_1)\cap\xi^{-1}(B_2))\Big)\mathbb{P}(A)=$$
$$=\mathbb{P}(\xi^{-1}(B_1)\cup\xi^{-1}(B_2))\mathbb{P}(A)=\mathbb{P}(\xi^{-1}(B_1\cup B_2))\mathbb{P}(A),$$

а це і означає знову, що $B_1\cup B_2\in\mathcal{M}_0$. Отже, за доведеним $\mathcal{M}_0$ — алгебра, яка містить всі скінченні об'єднання півінтервалів вигляду $[a,b)$, позаяк $[a,b)=(-\infty,b)\setminus(-\infty,a)$.
Якщо ми тепер доведемо, що $\mathcal{M}_0$ — монотонний клас множин, то за критерієм σ-алгебри це означатиме, що $\mathcal{M}_0$ — σ-алгебра.
Достатньо перевірити, що для довільної послідовності $(B_j),\ B_j\in\mathcal{M}_0$,

$$B_j\subset B_{j+1}\ (j\geq 1)\Longrightarrow B\stackrel{def}{=}\bigcup_{n=1}^{+\infty}B_n\in\mathcal{M}_0.$$

За неперервністю ймовірності знову маємо

$$\mathbb{P}(\xi^{-1}(B)\cap A) = \mathbb{P}\Big(\xi^{-1}\Big(\bigcup_{n=1}^{+\infty} B_n\Big)\cap A\Big) =$$

$$= \mathbb{P}\Big(\bigcup_{n=1}^{+\infty} \xi^{-1}\Big(B_n\Big)\cap A\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\xi^{-1}\Big(B_n\Big)\cap A\Big) =$$

$$= \mathbb{P}(A)\lim_{n\to+\infty} \mathbb{P}\Big(\xi^{-1}\Big(B_n\Big)\Big) = \mathbb{P}(A)\mathbb{P}\Big(\bigcup_{n=1}^{+\infty} \xi^{-1}\Big(B_n\Big)\Big) =$$

$$= \mathbb{P}(\xi^{-1}(B))\mathbb{P}(A),$$

тобто, $B \stackrel{def}{=} \bigcup\limits_{n=1}^{+\infty} B_n \in \mathcal{M}_0$. Отже, твердження доведене повністю.

Використовуючи тепер метод математичної індукції за кількістю випадкових величин, за допомогою цього твердження вже нескладно завершити доведення теореми.

Вправа. Завершити доведення теореми.

2.18. Випадкові вектори з незалежними компонентами

Розглянемо випадковий вектор $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$, тобто, вектор, компоненти якого ξ_j — випадкові величини.

Означення (функції спільного розподілу)

Функцією розподілу випадкового вектора $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ (функцією спільного розподілу випадкових величин $\xi_1, \xi_2, \ldots, \xi_n$) називається функція $F_\xi \colon \mathbb{R}^n \to [0, 1]$, яка для кожного $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ визначається формулою

$$F_\xi(x) = \mathbb{P}\{\omega \colon \xi_1(\omega) < x_1, \ldots, \xi_n(\omega) < x_n\} = \mathbb{P}\Big(\bigcap_{j=1}^{n} \{\omega \colon \xi_j(\omega) < x_j\}\Big).$$

Вправа. Довести, що:

1. $F_{(\xi_1,\xi_2,\ldots,\xi_n)}(+\infty, +\infty, \ldots, +\infty) = 1$.
2. $F_{(\xi_1,\xi_2,\ldots,\xi_n)}(+\infty, x_2, \ldots, x_n) = F_{(\xi_2,\ldots,\xi_n)}(x_2, \ldots, x_n)$.
3. $F_{(\xi_1,\xi_2,\ldots,\xi_n)}(-\infty, -\infty, \ldots, -\infty) = F_{(\xi_1,\xi_2,\ldots,\xi_n)}(-\infty, x_2, \ldots, x_n) = 0$.

Пригадаємо, що випадковий вектор $\xi = (\xi_1, \ldots, \xi_n)$ є вектором з незалежними компонентами, якщо $(\xi_j)_{j=1}^n$ — послідовність незалежних випадкових величин. Тоді критерій незалежності випадкових величин можна переписати у такому вигляді: випадковий вектор є вектором з незалежними компонентами тоді і тільки тоді, коли його функція розподілу є добутком функцій розподілу його компонент. Власне, маємо таку теорему.

Теорема (критерій незалежності випадкових величин)

Послідовність випадкових величин $(\xi_j)_{j=1}^n$ є послідовністю незалежних випадкових величин тоді і тільки тоді, коли

$$(\forall x \in \mathbb{R}^n)\colon \ F_\xi(x) = \prod_{j=1}^n F_{\xi_j}(x_j), \quad \text{де} \ \ \xi = (\xi_1, \ldots, \xi_n).$$

Теорема (незалежність борелевих функцій від випадкових величин)

1. Нехай $g\colon \mathbb{R}^n \to \mathbb{R}$, $h\colon \mathbb{R}^m \to \mathbb{R}$ — борелеві функції. Якщо $(\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_m)$ — вектор з незалежними компонентами, то випадкові величини $\xi = g(\xi_1, \ldots, \xi_n)$ і $\eta = h(\eta_1, \ldots, \eta_m)$ — незалежні.

2. Якщо $(\xi_n)_{n=1}^N$ — послідовність незалежних випадкових величин, а $(g_n)_{n=1}^N$ — борелеві функції на $\mathbb{R}$. Тоді $(\eta_n)_{n=1}^N$, $\eta_n = g_n(\xi_n)$, — послідовність незалежних випадкових величин.

Доведення Доведемо п.2 теореми при $N = 2$ (п.1 при $n = 1,\ m = 1$). Повне доведення обидвох пунктів теореми здійснюється за цією ж схемою. Отже, нехай ξ_1, ξ_2 — незалежні випадкові величини, а g_1, g_2 — борелеві функції на $\mathbb{R}$. Розглянемо довільні борелеві множини B_1, B_2 на дійсній прямій.

Оскільки

$$\eta_j^{-1}(B_j) = \{\omega\colon g_j(\xi_j(\omega)) \in B_j\} = \xi_j^{-1}(g_j^{-1}(B_j)),$$

а за означенням борелевої функції, $g_j^{-1}(B_j)$ — борелеві множини, то з незалежності випадкових величин ξ_j отримуємо

$$\mathbb{P}(\eta_1^{-1}(B_1) \cap \eta_2^{-1}(B_2)) = \mathbb{P}(\xi_1^{-1}(g_1^{-1}(B_1)))\mathbb{P}(\xi_2^{-1}(g_2^{-1}(B_2))) =$$
$$= \mathbb{P}(\eta_1^{-1}(B_1))\mathbb{P}(\eta_2^{-1}(B_2)),$$

звідки, η_1, η_2 — незалежні.

Для доведення п.2 в загальному, для будь-яких борелевих множин (B_j) на дійсній прямій, з незалежності (ξ_j), позаяк $g_j^{-1}(B_j)$ — борелеві множини, маємо

$$\mathbb{P}\Big(\bigcap_{j=1}^{N} \eta_j^{-1}(B_j)\Big) = \prod_{j=1}^{N} \mathbb{P}\Big(\xi_j^{-1}(g_j^{-1}(B_j))\Big) = \prod_{j=1}^{N} \mathbb{P}\Big(\eta_j^{-1}(B_j)\Big),$$

що й означає незалежність $(\eta_j)_{j=1}^{N}$.

Вправа. Довести п.1 теореми у повному обсязі.

Теорема

Якщо $(\xi_j)_{j=1}^k$ — послідовність незалежних випадкових величин, то

$$M\Big(\prod_{j=1}^k \xi_j\Big) = \prod_{j=1}^k M\xi_j,$$

зокрема, $(\xi_j)_{j=1}^k$ — послідовність попарно некорельованих випадкових величин.

Доведення Утворимо $\xi_j^+ = \sup\{\xi_j, 0\}$, $\xi_j^- = (-\xi_j)^+$; тоді $\xi_j = \xi_j^+ - \xi_j^-$, $\xi_j^\pm \geq 0$. Вище вже відзначалось, що якщо (ξ_j) — послідовність незалежних випадкових величин, то $(\xi_j^\pm)$ — також послідовність незалежних випадкових величин для будь-якого фіксованого вибору послідовності плюсів і мінусів у верхньому індексі.

Якщо твердження теореми буде доведено для невід'ємних випадкових величин, то з рівності

$$\prod_{j=1}^{k} \xi_j = \prod_{j=1}^{k} (\xi_j^+ - \xi_j^-) = \sum (-1)^s \xi_1^{\pm} \xi_2^{\pm} \dots \xi_k^{\pm},$$

де в останній сумі комбінації знаків $\pm$ у верхньому індексі є різними для різних доданків, а для кожного множника кожного разу вибрано лише один знак "$+$"або "$-$"; s дорівнює кількості знаків "$-$"у верхніх індексах в даному доданку. Тоді за незалежністю елементів послідовності $(\xi_j^{\pm})$

$$M\Big(\prod_{j=1}^{k} \xi_j\Big) = \sum (-1)^s M(\xi_1^{\pm} \xi_2^{\pm} \dots \xi_k^{\pm}) =$$

$$= \sum (-1)^s M\xi_1^{\pm} M\xi_2^{\pm} \dots M\xi_k^{\pm} = \prod_{j=1}^{k} (M\xi_j^+ - M\xi_j^-) = \prod_{j=1}^{k} M\xi_j.$$

Залишається розглянути випадок, коли (ξ_j) — послідовність невід'ємних незалежних випадкових величин. Для кожного $j,\ 1 \le j \le k$, побудуємо неспадні послідовності $(\xi_j^{(n)})_{n=1}^{+\infty}$ простих невід'ємних випадкових величин такі, що $\xi_j^{(n)} \to \xi_j (n \to +\infty)$ і для кожного n послідовність $(\xi_j^{(n)})_{j=1}^{k}$ є послідовністю незалежних випадкових величин:

1^0. $(\xi_j^{(n)})_{j=1}^{k}$ — послідовність невід'ємних незалежних випадкових величин для кожного фіксованого $n \ge 1$.

2^0. $(\forall j,\ 1 \le j \le k)(\forall \omega \in \Omega)(\forall n \ge 1)\colon \xi_j^{(n)}(\omega) \le \xi_j^{(n+1)}(\omega)$.

3^0. $\lim\limits_{n \to +\infty} \xi_j^{(n)}(\omega) = \xi_j(\omega)$.

Тоді з невід'ємності $(\xi_j^{(n)})$ і монотонності за n, випливає, що

$$\eta^{(n)} \stackrel{def}{=} \xi_1^{(n)}\xi_2^{(n)}\dots\xi_k^{(n)} \nearrow,\ \eta^{(n)} \geq 0\ (n \geq 1),\ \lim_{n\to+\infty}\eta^{(n)} = \xi_1\xi_2\dots\xi_k.$$

Вище було доведено (твердження 11), що для незалежних простих випадкових величин $(\xi_j^{(n)})_{j=1}^k$

$$M\eta^{(n)} = M\Big(\prod_{j=1}^k \xi_j^{(n)}\Big) = \prod_{j=1}^k M\xi_j^{(n)}$$

Граничний перехід при $n \to +\infty$ в останній рівності завершує справу. Перевіримо його законність, тобто обгрунтуємо, що $\lim\limits_{n\to+\infty} M\eta^{(n)} = M\lim\limits_{n\to+\infty}\eta^{(n)}$, а також, що $\lim\limits_{n\to+\infty} M\xi_j^{(n)} = M\lim\limits_{n\to+\infty}\xi_j^{(n)}$.

Справді, за означенням інтеграла $\lim\limits_{n\to+\infty} M\eta^{(n)} = M(\xi_1\xi_2\cdots\xi_k)$, але
$M\lim\limits_{n\to+\infty}\eta^{(n)} = M(\xi_1\xi_2\cdots\xi_k)$. Подібно міркуємо з іншою рівністю $\lim\limits_{n\to+\infty} M\xi_j^{(n)} = M\xi_j$ та $M\lim\limits_{n\to+\infty}\xi_j^{(n)} = M\xi_j$.

Залишається побудувати послідовності $(\xi_j^{(n)})$ з потрібними властивостями. Виберемо

$$\xi_j^{(n)}(\omega) \stackrel{def}{=} \begin{cases} \frac{i}{2^n}, & \frac{i}{2^n} \le \xi_j(\omega) < \frac{i+1}{2^n} \quad (0 \le i \le n \cdot 2^n - 1), \\ 0, & \xi_j(\omega) \ge n. \end{cases}$$

Вправа. Нехай (ξ_j) — прості випадкові величини, а $X^{(j)} = \{x_k^{(j)}\}$ їхні множини значень, відповідно. (ξ_j) — послідовність незалежних випадкових величин тоді і тільки тоді, коли

$$(\forall (k_j)) \colon \ \mathbb{P}\Big(\bigcap_{j=1}^{n} \xi_j^{-1}(\{x_{k_j}^{(j)}\})\Big) = \prod_{j=1}^{n} \mathbb{P}(\xi_j^{-1}(\{x_{k_j}^{(j)}\}).$$

Вказівка. В бік необхідності це звичайна інтерпретація означення незалежності. В бік достатності, зауважимо, що для кожної борелевої множини B_j виконується $\xi_j^{-1}(B_j) = \bigcup_{k\colon x_k^{(j)} \in B_j} \xi_j^{-1}(\{x_k^{(j)}\})$ і, отже,

$$\bigcap_{j=1}^{n} \xi_j^{-1}(B_j\{x_{k_j}^{(j)}\}) = \bigcup_{k_j\colon x_{k_j}^{(j)} \in B_j} \bigcap_{j=1}^{n} \xi_j^{-1}(\{x_{k_j}^{(j)}\}).$$

Залишається послідовно скористатись адитивністю ймовірності й умовою.

Для побудованої послідовності для всіх (s_j), $0 \le s_j \le n \cdot 2^n - 1$ маємо

$$\mathbb{P}\left(\bigcap_{j=1}^{k} (\xi_j^{(n)})^{-1}\left(\left\{\frac{s_j}{2^n}\right\}\right)\right) = \mathbb{P}\left(\bigcap_{j=1}^{k} \xi_j^{-1}\left(\left[\frac{s_j}{2^n}, \frac{s_j+1}{2^n}\right)\right)\right) =$$

$$= \prod_{j=1}^{k} \mathbb{P}\left(\xi_j^{-1}\left[\frac{s_j}{2^n}, \frac{s_j+1}{2^n}\right)\right) = \prod_{j=1}^{k} \mathbb{P}\left((\xi_j^{(n)})^{-1}\left(\left\{\frac{s_j}{2^n}\right\}\right)\right).$$

За вправою, це й означає потрібну незалежність. Перевіримо монотонність. Оскільки

$\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right] = \left[\frac{2i}{2^{n+1}}, \frac{2i+1}{2^{n+1}}\right) \bigcup \left[\frac{2i+1}{2^{n+1}}, \frac{2i+2}{2^{n+1}}\right)$, то

$$\xi_j(\omega) \in \left[\frac{2i}{2^{n+1}}, \frac{2i+1}{2^{n+1}}\right) \Longrightarrow \xi_j^{(n+1)}(\omega) = \frac{2i}{2^{n+1}} = \frac{i}{2^n} = \xi_j^{(n)}(\omega),$$

$$\xi_j(\omega) \in \left[\frac{2i+1}{2^{n+1}}, \frac{2i+2}{2^{n+1}}\right) \Longrightarrow \xi_j^{(n+1)}(\omega) = \frac{2i+1}{2^{n+1}} > \frac{2i}{2^{n+1}} = \frac{i}{2^n} = \xi_j^{(n)}(\omega).$$

Вправа. Переконатись, що $\lim\limits_{n\to+\infty} \xi_j^{(n)}(\omega) = \xi_j(\omega)$.

2.19. Збіжність майже напевно і збіжність за ймовірністю

Означення (збіжності майже напевно м.н.)

Послідовність випадкових величин (η_n) називається збіжною майже напевно (м.н) до випадкової величини η $\quad (\eta_n \to \eta$ м.н.), якщо

$$\mathbb{P}\{\omega : \lim_{n\to+\infty} \eta_n(\omega) = \eta(\omega)\} = 1.$$

З курсу "Теорії міри та інтегралу"добре відомо, що якщо послідовність вимірних функцій збіжна майже скрізь, то вона збіжна за мірою.

Твердження (15)

Якщо $\eta_n \to \eta$ (м.н.) $\Rightarrow \eta_n \xrightarrow{\mathbb{P}} \eta$.

Навпаки, взагалі кажучи, не правильно. Але правильне таке твердження.

Твердження (16)

Якщо $\eta_n \xrightarrow{\mathbb{P}} \eta$, то $(\exists n_k \uparrow +\infty) : \eta_{n_k} \to \eta$ (м.н.).

Для повноти викладу наводимо доведення тверджень 15 і 16.

Доведення твердження 15. Нехай $A \stackrel{def}{=} \{\omega : \lim\limits_{n\to+\infty} \eta_n(\omega) = \eta(\omega)\}$.

Спочатку зауважимо, що $\omega \in A \Leftrightarrow \lim_{n\to+\infty} \eta_n(\omega) = \eta(\omega) \Leftrightarrow$
$(\forall \varepsilon > 0)(\exists N)$
$(\forall n \geq N) : |\eta_n(\omega) - \eta(\omega)| < \varepsilon \Leftrightarrow (\forall k \in \mathbb{N})(\exists N)(\forall n \geq N) :$
$|\eta_k(\omega) - \eta(\omega)| < \frac{1}{k}$.

Позначимо тепер

$$A_{n,k} = \{\omega : |\eta_n(\omega) - \eta(\omega)| < \frac{1}{k}\}.$$

Тоді, з щойно записаного ланцюжка еквівалентностей випливає, що

$$A = \bigcap_{k=1}^{+\infty} \bigcup_{N=1}^{+\infty} \bigcap_{n=N}^{+\infty} A_{n,k}.$$

За умовою $\mathbb{P}(A) = 1$, $A = \bigcap\limits_{k=1}^{+\infty} A_k$, $A_k = \bigcup\limits_{N=1}^{+\infty} \bigcap\limits_{n=N}^{+\infty} A_{n,k}$.
Скористаємось тепер таким загальним твердженням.

Твердження (17)

$$\mathbb{P}(\bigcap_{k=1}^{n} A_k) = 1 \;\Leftrightarrow\; (\forall k,\ 1 \le k \le n)\colon\ \mathbb{P}(A_k) = 1.$$

Вправа. Довести твердження 17.

Позначимо $B_{N,k} \stackrel{def}{=} \bigcap\limits_{n=N}^{+\infty} A_{n,k}$. Оскільки $B_{N,k} \subset B_{N+1,k}$ $(N \ge 1)$ для кожного фіксованого k, то з твердження 17 за теоремою неперервності ймовірності випливає, що

$$\mathbb{P}(A) = 1 \quad \Longleftrightarrow \quad (\forall k \ge 1)\colon \quad \lim_{N\to+\infty} \mathbb{P}(B_{N,k}) = \mathbb{P}(A_k) = 1.$$

Тобто, ми встановили такий критерій збіжності м.н.

Твердження (критерій збіжності майже напевно)

Для послідовності випадкових величин (η_n) і випадкової величини η

$$\eta_n \to \eta \text{ м.н.} \iff (\forall k \geq 1)\colon \lim_{N\to+\infty} \mathbb{P}\Big(\bigcap_{n=N}^{\infty} \{|\eta_n(\omega) - \eta(\omega)| < \frac{1}{k}\}\Big) = 1$$

або у рівносильному вигляді

$$\eta_n \to \eta \text{ м.н.} \iff (\forall \varepsilon > 0)\colon \lim_{N\to+\infty} \mathbb{P}\Big(\bigcap_{n=N}^{\infty} \{|\eta_n(\omega) - \eta(\omega)| < \varepsilon\}\Big) = 1.$$

Оскільки $\bigcap\limits_{n=N}^{\infty} \{|\eta_n(\omega) - \eta(\omega)| < \varepsilon\} \subset \{|\eta_N(\omega) - \eta(\omega)| < \varepsilon\}$, то з критерію збіжності м.н. випливає, що

$$(\forall \varepsilon > 0)\colon \quad \lim_{N\to+\infty} \mathbb{P}(\{|\eta_N(\omega) - \eta(\omega)| < \varepsilon\}) = 1.$$

Іншими словами ми довели, що $\eta_n \xrightarrow{\mathrm{P}} \eta$ $(n \to +\infty)$.

З критерію збіжності м.н. отримуємо таку зручну у застосуваннях достатню умову збіжності м.н.

Твердження (достатня умова збіжності майже напевно)

Якщо для кожного $\varepsilon > 0$ збіжний ряд

$$\sum_{n=1}^{\infty} \mathbb{P}\{\omega\colon |\eta_n(\omega) - \eta(\omega)| \geq \varepsilon\} < +\infty$$

то $\eta_n \to \eta$ м.н.

Справді, для кожного $\varepsilon > 0$ за монотонністю ймовірності при $N \to +\infty$

$$\mathbb{P}(\bigcup_{n=N}^{\infty} \{\omega : |\eta_n(\omega) - \eta(\omega)| \geq \varepsilon\}) \leq \sum_{n=N}^{+\infty} \mathbb{P}\{\omega : |\eta_n(\omega) - \eta(\omega)| \geq \varepsilon\} \to 0$$

як залишок збіжного ряду. Отже, за критерієм збіжності майже напевно $\eta_n \to \eta$ м.н.

2.20. Лема Бореля-Кантеллі

Теорема (лема Бореля-Кантеллі)

Нехай (A_n) — послідовність подій.
1. Якщо $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < +\infty$, то серед подій (A_n) з ймовірністю, що дорівнює одиниці, відбувається скінченна кількість подій.

2. Якщо (A_n) — послідовність незалежних в сукупності подій таких, що з ймовірністю, яка дорівнює одиниці, серед (A_n) відбувається скінченна кількість подій, то $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < +\infty$.

Доведення. Через C позначимо подію "серед (A_n) відбувається нескінченна кількість подій". Зауважимо, що

$$C = \bigcap_{n=1}^{+\infty} \bigcup_{s=n}^{+\infty} A_s.$$

Справді, $\omega \in C \Leftrightarrow (\exists n_k \uparrow +\infty \ (k \to +\infty))\ (\forall k \geq 1)\colon\ \omega \in A_{n_k}$.

Останнє є рівносильним з тим, що $(\forall n \geq 1)\colon \omega \in \bigcup_{s=n}^{+\infty} A_s$, що виконується $\Leftrightarrow \omega \in \bigcap_{n=1}^{+\infty} \bigcup_{s=n}^{+\infty} A_s$.

Доведемо спочатку п.1. Для цього досить довести, що $\mathbb{P}(C) = 0$. Оскільки $C \subset \bigcup\limits_{s=N}^{\infty} A_s$, то, скориставшись монотонністю ймовірності і збіжністю ряду, отримаємо

$$0 \leq \mathbb{P}(C) \leq \sum_{s=N}^{\infty} \mathbb{P}(A_s) \to 0 \quad (n \to +\infty).$$

Отже, $\mathbb{P}(C) = 0$.

Доведемо тепер п.2. Оскільки, події (A_n) — незалежні в сукупності, то звідси випливає, що $(\overline{A}_n)$ також незалежні в сукупності.

Вправа. Довести цей факт.

За умовою $\mathbb{P}(C) = 0$. Розглянемо $B_n = \bigcup\limits_{s=n}^{+\infty} A_s$. Зауважимо, що $C = \bigcap\limits_{n=1}^{+\infty} B_n$, $B_n \supset B_{n+1}$ $(n \geq 1)$ — монотонно спадна послідовність подій.

Тому, за теоремою неперервності ймовірності

$$0 = \mathbb{P}(C) = \lim_{n\to+\infty} \mathbb{P}(B_n),$$

звідки за означенням границі на мові ε, вибираючи $\varepsilon = \frac{1}{2}$, зокрема, отримаємо $(\exists n_0)$: $\mathbb{P}(B_{n_0}) < \frac{1}{2}$. Звідси, $\mathbb{P}(\overline{B}_{n_0}) > \frac{1}{2}$. Але $\overline{B}_{n_0} = \bigcap\limits_{s=n_0}^{+\infty} A_s \subset \bigcap\limits_{s=n_0}^{m} A_s$ для всіх $m \geq n_0$. Тому, знову скориставшись монотонністю ймовірності, з незалежності подій $(\overline{A}_j)$ отримаємо

$$\frac{1}{2} < \mathbb{P}(B_{n_0}) \leq= \mathbb{P}(\bigcap_{s=n_0}^{m} \overline{A}_s) =$$

$$= \prod_{s=n_0}^{m} \mathbb{P}(\overline{A}_s) = \prod_{s=n_0}^{m} (1 - \mathbb{P}(A_s)) \leq \exp\Big\{ - \sum_{s=n_0}^{m} \mathbb{P}(A_s)\Big\},$$

при цьому ми скористались нерівністю $1 - x \leq e^{-x}$ $(x \in \mathbb{R})$.

Логарифмування початку і кінця цього ланцюжка перетворень, дає нерівність

$$\sum_{s=n_0}^{m} \mathbb{P}(A_s) < \ln 2 \quad (m \geq n_0).$$

Останнє означає, що часткові суми ряду $\sum_{s=n_0}^{\infty} \mathbb{P}(A_s)$ з невід'ємними членами обмежені в сукупності, звідки й отримуємо збіжність цього ряду.

Доведення твердження 16. Запишемо означення границі $\eta_n \xrightarrow{\mathbb{P}} \eta$ на мові ε

$$(\forall \varepsilon > 0)(\forall \varepsilon_1 > 0)(\exists n_0)(\forall n > n_0(\varepsilon_1))\colon\ \mathbb{P}\{\omega\colon |\eta_n(\omega) - \eta(\omega)| \geq \varepsilon\} < \varepsilon_1.$$

Вибираючи для кожного $k \in \mathbb{N}$: $\varepsilon = \frac{1}{k}$, $\varepsilon_1 = \frac{1}{k^2}$, звідси для $A_{n_k} = \{\omega\colon |\eta_{n_k}(\omega) - \eta(\omega)| \geq \frac{1}{k}\}$, $n_k = n_0(\varepsilon_1)$, отримаємо $\mathbb{P}(A_{n_k}) < \frac{1}{k^2}$, тобто, ряд $\sum_{k=1}^{\infty} \mathbb{P}(A_{n_k})$ — збіжний.

За лемою Бореля-Кантеллі, серед елементів послідовності подій (A_{n_k}) з ймовірністю, що дорівнює 1, відбувається скінченна кількість подій. Тому, існує $A \in \mathcal{A}$, $\mathbb{P}(A) = 1$ $\Big($за доведенням леми Бореля-Кантеллі $\overline{A} = \bigcap\limits_{k=1}^{+\infty} \bigcup\limits_{s=k}^{+\infty} A_{n_s}$, а, отже $A = \bigcup\limits_{k=1}^{+\infty} \bigcap\limits_{s=k}^{+\infty} \overline{A}_{n_s}\Big)$ така, що для всіх $\omega \in A$

$$(\exists k_0 = k_0(\omega))(\forall k \geq k_0)\colon\ \omega \in \overline{A}_{n_k}.$$

Тому з для всіх $\omega \in A$ і для всіх $k \geq k_0$: $|\eta_{n_k}(\omega) - \eta(\omega)| < \frac{1}{k}$, тобто,

$$\eta_{n_k}(\omega) \to \eta(\omega)\ (k \to +\infty).$$

Залишається пригадати, що $\mathbb{P}(A) = 1$.

2.21. Посилений закон великих чисел (ПЗВЧ)

Подальша наша мета — доведення посиленого закону великих чисел (ПЗВЧ). Власне, наприклад, виявиться, що за умов теореми Чебишова у випадку послідовності незалежних випадкових величин збіжність до нуля різниці середніх арифметичних випадкових величин і їхніх математичних сподівань є збіжністю майже напевно і, отже, сильнішою, ніж у теоремі Чебишова.

Теорема (Колмогорова)

Нехай (ξ_n) — послідовність незалежних випадкових величин зі скінченними дисперсіями $D\xi_n < \infty$ $(n \geq 1)$, таких, що

$$\sum_{n=1}^{+\infty} \frac{D\xi_n}{n^2} < +\infty.$$

Тоді

$$\frac{1}{n} \sum_{k=1}^{n} (\xi_k - M\xi_k) \to 0 \quad \text{м.н.}$$

Зауваження. У теоремі Чебишова ми припускали, що всі дисперсії обмежені в сукупності: $D\xi_n \leq c < +\infty$ $(n \geq 1)$. З цього, очевидно випливає, що $\sum_{n=1}^{\infty} \frac{D\xi_n}{n^2} < +\infty$.
У випадку, коли ξ_n — однаково розподілені випадкові величини, то у попередньому твердженні можна відмовитись навіть від умови існування дисперсій. Власне, це є змістом наступної теореми.

Теорема (Колмогорова-Хінчіна)

Якщо (ξ_j) — послідовність незалежних випадкових величин з однаковим розподілом $(F_{\xi_j} = F$ $(j \geq 1))$ і $\exists M\xi_j = a \neq \infty$ $(j \geq 1)$, то

$$\frac{1}{n}\sum_{j=1}^{n} \xi_j \to a \quad \text{м.н.} \ .$$

Для доведення теореми Колмогорова нам потрібна така лема.

Лема (нерівність Колмогорова)

Нехай (ξ_n) — незалежні випадкові величини, $\eta_n = \sum_{k=1}^{n} \xi_k$.
Якщо $M\xi_n = 0$ $(n \geq 1)$ і $D\xi_n < +\infty (n \geq 1)$, то

$$(\forall N \geq 1)\ (\forall a > 0) : \mathbb{P}\{\omega : \sup_{1 \leq n \leq N} |\eta_n| \geq a\} \leq \frac{1}{a^2} \sum_{k=1}^{N} D\xi_k.$$

Доведення нерівності Колмогорова. Для $j \in \mathbb{N}$ приймемо
$A_j = \{\omega : |\eta_j| < a\}$; $B_1 = \overline{A}_1$, $B_2 = A_1 \cap \overline{A}_2, \ldots,$
$B_j = A_1 \cap A_2 \cap \ldots \cap A_{j-1} \cap \overline{A}_j$; $B = \bigcup\limits_{j=1}^{N} B_j$.

1. Доведемо спочатку, що
$A \stackrel{def}{=} \{\omega : \sup\{|\eta_n(\omega)|:\ 1 \leq n \leq N\} \geq a\} = B.$

$\omega \in A \Leftrightarrow \sup_{1\le n\le N} |\eta_n(\omega)| \ge a \Leftrightarrow (\exists n = n(\omega)) : |\eta_n(\omega)| \ge a \Leftrightarrow$
$\omega \in \overline{A}_n, \omega \in A_k, 1 \le k \le n-1 \Leftrightarrow \omega \in B_n; |\eta_k(\omega)| < a,$
$1 \le k \le n-1; \Leftrightarrow \omega \in B.$
$M\xi_k = 0$, $A = \{\omega : \sup_{1\le k\le N} |\eta_n(\omega)| \ge a\}$, $N < +\infty$; $\eta_n = \sum_{k=1}^n \xi_k$.
2. Доведемо, що $B_k \cap B_j = \varnothing (k \ne j)$. Для визначеності виберемо $k > j$. Тоді, якщо $\omega \in B_k \cap B_j$, то $\omega \in \overline{A}_j$ і $\omega \in A_j$, тобто, $\omega \in \overline{A}_j \cap A_j = \varnothing$.
З попарної несумісності елементів послідовності (B_j) негайно отримуємо таку рівність для характеристичних функцій

$$\chi_B = \sum_{n=1}^N \chi_{B_n}.$$

3. Перевіримо тепер, що $\sum_{k=1}^N D\xi_k = M\eta_N^2$. Маємо

$$\eta_N^2 = \left(\sum_{k=1}^N \xi_n\right)^2 = \sum_{n=1}^N \xi_n^2 + 2\sum_{1\le k<n\le N} \xi_k\xi_n.$$

Пригадаємо, що з незалежності ξ_k, ξ_n $(n \ne k)$ випливає, що

$$M(\xi_k\xi_n) = M\xi_k M\xi_n = 0.$$

Звідки,

$$M\eta_N^2 = \sum_{n=1}^N M\xi_n^2 + 2\sum_{1\le k<n\le N} M(\xi_k\xi_n) = \sum_{n=1}^N M\xi_n^2 = \sum_{n=1}^N D\xi_n,$$

позаяк, $D\xi_n = M\xi_n^2 - (M\xi_n)^2 = M(\xi_n)^2$

4. Оскільки

$$\sum_{k=1}^N D\xi_k = M\eta_N^2 = \int_\Omega \eta_N^2 \mathbb{P}(d\omega) \ge \int_\Omega \eta_N^2 \chi_B \mathbb{P}(d\omega),$$

то за п.2, записуючи

$\eta_N^2 = \eta_k^2 + (\eta_N - \eta_k)^2 + 2\eta_k(\eta_N - \eta_k) \ge \eta_k^2 + 2\eta_k(\eta_N - \eta_k),$

отримуємо

$$\sum_{k=1}^N D\xi_k \ge \sum_{k=1}^n \int_\Omega \eta_N^2 \chi_{B_k} \mathbb{P}(d\omega) \ge$$

$$\ge \sum_{k=1}^N \left(\int_\Omega \eta_k^2 \chi_{B_k} \mathbb{P}(d\omega) + +2 \int_\Omega (\eta_k \chi_{B_k})(\eta_N - \eta_k) \mathbb{P}(d\omega) \right) =$$

$$= \sum_{k=1}^N \left(\int_{B_k} \eta_k^2 \mathbb{P}(d\omega) + 2M((\chi_{B_k}\eta_k)(\eta_N - \eta_k)) \right)$$

Для кожного k, $1 \le k \le N$ визначимо

$$G_k = \Big\{(x_1, \dots, x_k) \in \mathbb{R}^k \colon |x_1| < a, |x_1 + x_2| < a, \dots,$$

$$\dots, \Big|\sum_{s=1}^{k-1} x_s\Big| < a, \Big|\sum_{s=1}^{k} x_s\Big| \ge a\Big\},$$

і розглянемо функцію $g_k \colon \mathbb{R}^k \to \mathbb{R}, \quad g_k(x) = \chi_{G_k}(x) \sum_{s=1}^{k} x_s$, де χ_{G_k} — індикатор множини G_k. Зрозуміло, що $g_k(x)$ борелева функція на $\mathbb{R}^k$, а також, що $\chi_{B_k}(\omega)\eta_k(\omega) = g_k(\xi_1, \xi_2, \dots, \xi_k)$. Тому з теореми про незалежність випадкових величин, утворених суперпозиціями борелевих функцій від незалежних випадкових величин $\Big($ Використовуємо теорему: якщо є дві борелеві функції $g(x_1, \dots, x_n)$, $h(x_1, \dots, x_m)$, а $(\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_m)$ — вектор з незалежними компонентами, то $g = g(\xi_1, \dots, \xi_n)$ і $h = h(\eta_1, \dots, \eta_m)$ — незалежні випадкові величини$\Big)$, і з незалежності (ξ_j) випливає, що для кожного $j \ge k+1$ випадкові величини ξ_j і $\chi_{B_k}(\omega)\eta_k(\omega) = g_k(\xi_1, \xi_2, \dots, \xi_k)$ — незалежні.

Тому некорельовані, звідки $M(\xi_j \cdot (\chi_{B_k}\eta_k)) = M\xi_j \cdot M(\chi_{B_k}\eta_k) = 0$. Звідси,

$$M((\chi_{B_k}\eta_k)(\eta_N - \eta_k)) = \sum_{j=k+1}^{N} M((\chi_{B_k}\eta_k)\xi_j) = 0.$$

Отже, пригадуючи означення B_k, остаточно отримуємо

$$\sum_{k=1}^{N} D\xi_k \geq$$

$$\geq \sum_{k=1}^{N} \int_{B_k} \eta_k^2 \mathbb{P}(d\omega) \geq a^2 \sum_{k=1}^{N} \mathbb{P}(B_k) = a^2\mathbb{P}(B) = a^2\mathbb{P}(A).$$

Доведення теореми Колмогорова.

Оскільки для послідовності $\xi'_k = \xi_k - M\xi_k$ маємо $D\xi'_k = D\xi_k$, $M\xi'_k = 0$, $\frac{1}{n}\sum_{k=1}^{n}(\xi_k - M\xi_n) = \frac{1}{n}\sum_{k=1}^{n}\xi'_k$, то теорему Колмогорова досить довести для послідовності (ξ_k) з нульовими математичними сподіваннями $M\xi_k = 0$. При цьому доводитимемо, що $\frac{1}{n}\sum_{k=1}^{n}\xi_k \to 0$ м.н.

Позначимо $\zeta_n = \sup\{\left|\sum_{k=1}^{s}\xi_k\right| : s \le 2^n\}$. Тоді, для $2^{n-1} \le s < 2^n$

$$\frac{1}{s}\left|\sum_{k=1}^{s}\xi_k\right| \le \frac{\zeta_n}{2^{n-1}}.$$

Достатньо довести, що $\frac{\zeta_n}{2^n} \to 0$ (м.н.). Для цього перевіримо виконання достатньої умови збіжності майже напевно

$$(\forall\varepsilon > 0)\colon \ \sum_{n=1}^{+\infty}\mathbb{P}\{\omega : |\theta_n(\omega)| \ge \varepsilon\} < +\infty \Longrightarrow \ \theta_n \to 0 \ \text{ м.н.}$$

Відповідно до цього, досить довести, що для кожного $\varepsilon > 0$: $\sum_{n=1}^{+\infty}\mathbb{P}\{\omega : \frac{\zeta_n}{2^n} \ge \varepsilon\} < +\infty$. За нерівністю Колмогорова можемо написати, що

$$(\forall a > 0) : \mathbb{P}\{\omega : \zeta_n > a\,2^n\} \le \frac{1}{a^2 2^{2n}}\sum_{k=1}^{2^n} D\xi_k.$$

Отже,

$$\mathbb{P}\{\omega : \zeta_n > a\,2^n\} \leq \frac{1}{a^2 2^{2n}} \sum_{k=1}^{2^n} D\xi_k.$$

Нехай $m_k = \min\{n : 2^n \geq k\}$. Оскільки, $2^{m_k} \geq k$ і, тому, $\left(\frac{1}{4}\right)^{m_k} \leq \frac{1}{k^2}$, то

$$\sum_{2^n \geq k} \frac{1}{4^n} = \sum_{n=m_k} \frac{1}{4^n} = \frac{\left(\frac{1}{4}\right)^{m_k}}{1 - \frac{1}{4}} = \frac{4}{3}\left(\frac{1}{4}\right)^{m_k} \leq \frac{4}{3}\frac{1}{k^2},$$

Тоді, для кожного $\varepsilon > 0$

$$\sum_{n=1}^{+\infty} \mathbb{P}\left\{\omega : \frac{\zeta_n}{2^n} > \varepsilon\right\} = \sum_{n=1}^{+\infty} \frac{1}{\varepsilon^2 2^{2n}} \sum_{k=1}^{2^n} D\xi_k = \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} D\xi_k \sum_{k \leq 2^n} \frac{1}{2^{2n}} \leq$$

$$\leq \frac{4}{3\varepsilon^2} \sum_{k=1}^{\infty} \frac{D\xi_k}{k^2}.$$

Але останній ряд збіжний за умовою.

Доведення теореми Колмогорова-Хінчіна. Для $n \in \mathbb{N}$ приймемо

$$A_n = \{\omega : |\xi_n(\omega)| \le n\},$$

і визначимо такі випадкові величини $\xi'_n(\omega) = \begin{cases} \xi_n(\omega), & \omega \in A_n, \\ 0, & \omega \notin A_n, \end{cases}$

$\xi''_n = \xi_n - \xi'_n$. Очевидно, що $\xi_n = \xi'_n + \xi''_n$. Наше доведення складається з декількох частин. Спочатку за теоремою Колмогорова доведемо, що

1. $\dfrac{1}{n}\sum_{k=1}^{n} \xi'_k \to a$ м.н.,

а потім за лемою Бореля-Кантелі доведемо, що

2. $\dfrac{1}{n}\sum_{k=1}^{n} \xi''_k \to 0$ м.н.

1. Перевіримо, що для послідовності (ξ'_k) виконуються умови теореми Колмогорова. Доведемо, що

$$\sum_{k=1}^{+\infty} \frac{D\xi'_k}{k^2} < +\infty.$$

Оскільки,

$$D\xi'_k = M(\xi'_k)^2 - (M\xi'_k)^2 \le M(\xi'_k)^2 = \int_\Omega (\xi'_k)^2 \mathbb{P}(d\omega) = \int_{A_k} (\xi_k)^2 \mathbb{P}(d\omega) =$$

$$= \int_{|x|\le k} x^2 dF(x) = \sum\nolimits_{s=1}^{k} \int_{s-1<|x|\le s} x^2 dF(x),$$

то $\sum\nolimits_{k=1}^{+\infty} \dfrac{D\xi'_k}{k^2} = \sum\nolimits_{k=1}^{+\infty} \dfrac{1}{k^2} \left(\sum\nolimits_{s=1}^{k} \int_{s-1<|x|\le s} x^2 dF(x) \right) =$

$$= \sum\nolimits_{s=1}^{+\infty} \left(\sum\nolimits_{k=1}^{+\infty} \frac{1}{k^2} \int_{s-1<|x|\le s} x^2 dF(x) \right) \le$$

$$\le 2 \sum\nolimits_{s=1}^{+\infty} \left(\sum\nolimits_{k=s}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+1} \right) \right) \int_{s-1<|x|\le s} x^2 dF(x) \le$$

$$\le 2 \sum\nolimits_{s=1}^{+\infty} \frac{1}{s} \int_{s-1<|x|\le s} x^2 dF(x) \le 2 \sum\nolimits_{s=1}^{+\infty} \int_{s-1<|x|\le s} |x| dF(x) =$$

$$= 2 \int_{-\infty}^{+\infty} |x| dF(x) = 2M|\xi_1| < +\infty.$$

При цьому ми врахували, що

$$\frac{1}{k^2} \le \frac{2}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1} \quad \text{і} \quad \sum_{k=s}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+1} \right) = \frac{1}{s}.$$

Зауважимо тепер, що

$M\xi'_k = \int_\Omega \xi'_k \mathbb{P}(d\omega) = \int_{A_k} \xi_k \mathbb{P}(d\omega) = \int_{-k}^{k} x dF(x)$, тому

$\lim\limits_{k\to+\infty} M\xi'_k = M\xi_1 = a$, звідки випливає, що

$\lim\limits_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} M\xi'_k = a.$

Отже, м.н., тобто, для всіх $\omega \in C_1$, $\mathbb{P}(C_1) = 1$, при $n \to +\infty$

$$\frac{1}{n}\sum_{k=1}^{n} \xi'_k = \frac{1}{n}\sum_{k=1}^{n}(\xi'_k - M\xi'_k) + \frac{1}{n}\sum_{k=1}^{n}(M\xi'_k - a) + a \to a.$$

Доведемо п.2. Нехай $B_n = \{\omega : \xi''_n(\omega) \neq 0\}$. Покажемо, що $B_n = \overline{A}_n$. Справді, $\omega \in \overline{A}_n \Leftrightarrow |\xi_n(\omega)| > n \Leftrightarrow \xi''_n(\omega) = \xi_n(\omega) \neq 0$ $\Leftrightarrow \omega \in B_n$.

Доведемо, що $\sum_{k=1}^{+\infty} \mathbb{P}(\overline{A}_k) = \sum_{k=1}^{+\infty} \mathbb{P}\{\omega : \xi''_k \neq 0\} < +\infty.$

Справді, оскільки $\mathbb{P}(\overline{A}_k) = \int_{|x|>k} dF(x)$, то

$$\sum_{k=1}^{+\infty} \mathbb{P}(\overline{A}_k) = \sum_{k=1}^{+\infty}\sum_{s=k}^{+\infty} \int_{s<|x|\le s+1} dF(x) = \sum_{s=1}^{+\infty}\left(\sum_{k=1}^{s} 1\right) \int_{s<|x|\le s+1} dF(x) =$$

$$= \sum_{s=1}^{+\infty} s \int_{s<|x|\le s+1} dF(x) \le \sum_{s=1}^{+\infty} \int_{s<|x|\le s+1} |x| dF(x) =$$

$$= \int_{1<|x|<+\infty} |x| dF(x) \le \int_{-\infty}^{+\infty} |x| dF(x) = M|\xi_1| < +\infty.$$

Ряд

$$\sum_{k=1}^{+\infty} \mathbb{P}\{\omega : \xi_k''(\omega) = 0\} < +\infty,$$

тому за лемою Бореля-Кантеллі серед подій $\overline{A}_k = \{\omega : \xi_k''(\omega) \ne 0\}$ з ймовірністю 1, відбувається скінченна кількість подій, звідки, існує $C_2 \in \mathcal{A}$ така, що $\mathbb{P}(C_2) = 1$ і для всіх $\omega \in C_2$ існує $k_0 = k_0(\omega)$ таке, що для всіх $k \ge k_0$: $\xi''(\omega) = 0$.

Звідси випливає, що для всіх $\omega \in C_2$

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k'' = \frac{1}{n}\sum_{k=1}^{k_0-1}\xi_k'' \to 0 \quad (n \to +\infty).$$

Тоді, для всіх $\omega \in C_1 \cap C_2$, при $n \to +\infty$

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k = \frac{1}{n}\sum_{k=1}^{n}\xi_k' + \frac{1}{n}\sum_{k=1}^{n}\xi_k'' \to a.$$

Залишається зауважити, що $\mathbb{P}(C_1 \cap C_2) = 1$. Теорема Колмогорова-Хінчіна повністю доведена.

2.22. Згортка розподілів випадкових величин

Згорткою розподілів F_1 і F_2 називається функція $F_1 * F_2 \colon \mathbb{R} \to [0,1]$, визначена для всіх $x \in \mathbb{R}$ за формулою

$$(F_1 * F_2)(x) \stackrel{def}{=} \int_{-\infty}^{+\infty} F_1(x-y)dF_2(y).$$

Вправа. Переконайтесь, що $F_1 * F_2$ — функція розподілу.

Теорема (теорема про згортку розподілів)

Нехай ξ_1 і ξ_2 — незалежні випадкові величини. Тоді

$$(\forall x \in \mathbb{R})\colon\ F_{\xi_1+\xi_2}(x) = (F_1 * F_2)(x) = (F_2 * F_1)(x),$$

де $F_j = F_{\xi_j}$ $(j \in \{1,2\})$.

Доведення. Нехай $F = F_{\xi_1+\xi_2}$. Тоді за формулою заміни змінної інтегрування

$$F(x) = \mathbb{P}\{\omega\colon \xi_1+\xi_2 < x\} = \int\limits_{\{\omega\colon \xi_1+\xi_2<x\}} \mathbb{P}(d\omega) = \iint\limits_{\{(t_1,t_2)\colon t_1+t_2<x\}} dF(t_1,t_2),$$

де $dF(t_1,t_2) = \nu_F(dt_1,dt_2)$ — міра Лебега-Стілт'єса, породжена функцією спільного розподілу випадкових величин ξ_1, ξ_2.

Оскільки, за означенням ν_F і за критерієм незалежності випадкових величин, застосованим до ξ_1, ξ_2

$$\nu_F([a,b)\times[c,d)) = F(b,d)+F(a,c)-F(b,c)-F(a,d) =$$
$$= F_1(b)F_2(d)+F_1(a)F_2(c)-F_1(b)F_2(c)-F_1(a)F_2(d) =$$
$$= (F_1(b)-F_1(a))(F_2(d)-F_2(c)) =$$
$$= \nu_{F_1}([a,b))\nu_{F_2}([c,d)) \stackrel{def}{=} \nu_{F_1,F_2}([a,b)\times[c,d)),$$

то нескладно тепер переконатись, що природно означені на алгебрі $\mathcal{A}_0$, породженій скінченними об'єднаннями прямокутників $\Pi = ([a,b)\times[c,d))$, міри ν_F та ν_{F_1,F_2}, за теоремою Каратеодорі-Лебега мають одне і те ж продовження на мінімальну σ-алгебру породжену алгеброю $\mathcal{A}_0$. Але, з одного боку, продовженням ν_F є міра Лебега-Стілт'єса, а з іншого, продовженням ν_{F_1,F_2} є прямий добуток мір $\nu_{F_1}\otimes\nu_{F_2} = dF_1\otimes dF_2$. Тому за теоремою Фубіні

$$F(x) = \iint_{\{(t_1,t_2):\ t_1+t_2<x\}} dF(t_1,t_2) = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{x-t_2} dF_1(t_1)\right)dF_2(t_2) =$$
$$= \int_{-\infty}^{+\infty} F_1(x-t_2)dF_2(t_2) = (F_1 * F_2)(x).$$

І, нарешті, з очевидної рівності $F_{\xi_1+\xi_2} = F_{\xi_2+\xi_1}$ випливає, що

$$F_1 * F_2 = F_2 * F_1.$$

Якщо функція розподілу F_2 має щільність f_2, то з теореми про формулу згортки негайно отримуємо такий наслідок.

Наслідок

Нехай ξ_1 і ξ_2 — незалежні випадкові величини, а f_2 — щільність розподілу ξ_2. Тоді

$$(\forall x \in \mathbb{R}): \ F_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} F_1(x-y) f_2(y) dy.$$

Наслідок

Нехай ξ_1 і ξ_2 — незалежні випадкові величини, а f_1 — щільність розподілу ξ_1. Тоді, сума випадкових величин $\xi_1 + \xi_2$ має щільність $f_{\xi_1+\xi_2}$ і м.с. за мірою Лебега

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} f_1(x-y) dF_2(y).$$

Доведення. Розглянемо функцію

$$g(x) \stackrel{def}{=} \int_{-\infty}^{+\infty} f_1(x-y)dF_2(y) \quad (x \in \mathbb{R}).$$

Для завершення доведення досить перевірити, що для кожного $x \in \mathbb{R}$

$$F(x) = \int_{-\infty}^{x} g(t)dt.$$

Звідси, як вже відзначалось вище, випливатиме, що $g(x) = f_{\xi_1+\xi_2}(x)$ м.с. за мірою Лебега. Отже, застосовуючи теорему Фубіні, послідовно отримуємо

$$\int_{-\infty}^{x} g(t)dt = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_1(t-y)dF_2(y)dt =$$

$$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{x} f_1(t-y)dt\right)dF_2(y) = \int_{-\infty}^{+\infty} F_1(x-y)dF_2(y) = F(x).$$

При цьому останню рівність ми записали за формулою згортки розподілів.

З попередніх наслідків негайно маємо ще такий наслідок про згортку щільностей.

Наслідок (7, згортка щільностей)

Нехай f_1 і f_2 — щільності розподілів випадкових величин ξ_1 і ξ_2. Тоді для щільності розподілу суми $\xi_1 + \xi_2$ м.с. за мірою Лебега виконується рівність

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} f_1(x-y) f_2(y) dy.$$

Згорткою щільностей f_1 та f_2 називають

$$(f_1 * f_2)(x) \stackrel{def}{=} \int_{-\infty}^{+\infty} f_1(x-y) f_2(y) dy.$$

Приклад (Згортка рівномірних розподілів). Розглянемо незалежні випадкові величини ξ і η, що мають рівномірний розподіл на проміжку $[a, b]$ $f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a, b), \\ 0, & x \notin [a, b]. \end{cases}$

З незалежності ξ і η за наслідком 7 про згортку щільностей отримуємо

$$f_{\xi+\eta}(x) = (f * f)(x) = \int_{-\infty}^{+\infty} f(x-y)f(y)dy = \frac{1}{b-a}\int_a^b f(x-y)dy =$$

$$= \left[\begin{array}{l} x-y=t \\ dy=-dt \end{array}\right] = \frac{1}{b-a}\int_{x-b}^{x-a} f(t)dt = \begin{cases} 0, & x<2a, \\ \frac{x-2a}{(b-a)^2}, & 2a<x<a+b, \\ \frac{2b-x}{(b-a)^2}, & a+b<x<2b, \\ 0, & x>2b. \end{cases}$$

Подібно знаходимо згортку незалежних рівномірних розподілів на різних проміжках.

2.23. Гамма-розподіл

Гамма-розподіл $\Gamma(\alpha, \beta)$ з параметрами $\alpha > 0, \beta > 0$ — це розподіл випадкової величини з такою щільністю

$$f(x) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \cdot e^{-\beta x}, & x > 0, \\ 0, & x < 0, \end{cases}$$

де $\Gamma(\alpha)$ — Γ-функція Ойлера $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$.

Вправа. Перевірити, що це справді щільність розподілу, тобто, що $f(x) \geq 0$ м.с. і $\int_0^{+\infty} f(x)dx = 1$.

Розподіл $\exp(\beta) \stackrel{def}{=} \Gamma(1, \beta)$ називається показниковим (експоненційним) розподілом, тобто, це розподіл зі щільністю

$$f(x) = \beta e^{-\beta x} \ (x > 0), \ f(x) = 0 \ (x < 0).$$

Твердження (теорема додавання Γ-розподілів)

Нехай ξ_1, ξ_2 — незалежні випадкові величини, $\xi_j \in \Gamma(\alpha_j, \beta)$. Тоді $\xi = \xi_1 + \xi_2 \in \Gamma(\alpha_1 + \alpha_2, \beta)$.

Доведення. За формулою згортки для щільностей

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} f_1(x-y) f_2(y) dy =$$

$$= \int_0^x f_1(x-y) f_2(y) dy = \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^x e^{-\beta(x-y)-\beta y}(x-y)^{\alpha_1-1} y^{\alpha_2-1} dy =$$

$$= \frac{\beta^{\alpha_1+\alpha_2} e^{-\beta x}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int\limits_0^x (x-y)^{\alpha_1-1} y^{\alpha_2-1} dy = \Big[y = tx,\ dy = x dt \Big] =$$

$$= \frac{\beta^{\alpha_1+\alpha_2} x^{\alpha_1+\alpha_2-1} e^{-\beta x}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \cdot B(\alpha_1, \alpha_2),$$

де $B(\alpha_1, \alpha_2) = \int_0^1 t^{\alpha_1-1}(1-t)^{\alpha_2-1} dt$ — β-функція Ойлера.

Пригадуючи, що $B(\alpha_1, \alpha_2) = \dfrac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$, вже елементарно завершуємо доведення.

Приклад (Розподіл Ерланга). Розподіл Ерланга — це розподіл випадкової величини

$$\eta = \xi_1 + \xi_2 + \ldots + \xi_n,$$

де (ξ_j) — незалежні випадкові величини, що мають один і той же показниковий розподіл $\xi_j \in \exp(\beta) \stackrel{def}{=} \Gamma(1, \beta)$, $\beta > 0$. Тоді за теоремою додавання розподіл Ерланга, це Γ-розподіл $\Gamma(n, \beta)$.

Вправа (згортка показникових розподілів). Нехай ξ_1, ξ_2– незалежні випадкові величини з показниковими розподілами $\xi_j \in \exp(\beta_j)$, $0 < \beta_1 < \beta_2$. Довести, що щільність розподілу суми $\xi_1 + \xi_2$ дорівнює

$$f(x) = \beta_1 \beta_2 \frac{e^{-\beta_1 x} - e^{-\beta_2 x}}{\beta_2 - \beta_1} \quad (x > 0), \quad f(x) = 0 \quad (x < 0).$$

2.24. Розподіл χ^2

Розподіл χ^2 з n-ступенями свободи — це розподіл такої випадкової величини

$$\chi^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2,$$

де (ξ_j) — незалежні випадкові величини, $\xi_j \in \mathcal{N}(0,1)$ $(1 \le j \le n)$.

Твердження

Розподіл χ^2 — це Γ-розподіл $\Gamma(\frac{n}{2}, \frac{1}{2})$.

Доведення. Досить перевірити, що якщо ξ є $\mathcal{N}(0,1)$, то $\xi^2 \in \Gamma(\frac{1}{2}, \frac{1}{2})$. Справді, $F(x) = 0$ для $x \le 0$, а для $x > 0$

$$F_{\xi^2}(x) = \mathbb{P}\{\omega\colon \xi^2(\omega) < x\} = \mathbb{P}\{-\sqrt{x} < \xi < \sqrt{x}\} = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}),$$

звідки, для $x > 0$

$$f_{\xi^2}(x) = F'_{\xi^2}(x) = \frac{1}{2\sqrt{x}}\Phi'(\sqrt{x}) + \frac{1}{2\sqrt{x}}\Phi'(-\sqrt{x}) = \frac{x^{-1/2}}{\sqrt{2\pi}}e^{-\frac{x}{2}}.$$

Звідси, $\xi^2 \in \Gamma(\frac{1}{2}, \frac{1}{2})$. Застосування теореми додавання для Γ-розподілів завершує доведення.

Розподіл $\eta = \xi_1^2 + \xi_2^2 + \xi_3^2$, де $\xi_j \in \mathcal{N}(0,1)$ і незалежні, у фізиці називають розподілом Максвела. З щойно доведеного випливає, що розподіл Максвела це розподіл $\Gamma(\frac{3}{2}, \frac{1}{2})$.

2.25. Згортка дискретних розподілів

Нехай $\xi_1,\ \xi_2$ — незалежні дискретні цілочисельні випадкові величини з розподілами $p_j(n) = \mathbb{P}\{\omega\colon \xi_j(\omega) = n\}$ $(n \in \mathbb{Z})$, $j \in \{1, 2\}$.

Пригадаємо спочатку, що для довільної випадкової величини η

$$\mathbb{P}\{\omega\colon \eta(\omega) = a\} = F_\eta(a+0) - F_\eta(a),$$

а за теоремою про згортку

$$F_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} F_1(x-y)dF_2(y) =$$

$$= \sum\nolimits_{m=-\infty}^{+\infty} F_1(x-m)(F_2(m+0) - F_2(m)) = \sum\nolimits_{m=-\infty}^{+\infty} F_1(x-m)p_2(m).$$

Тому, для кожного $n \in \mathbb{Z}$

$$p(n) \stackrel{def}{=} \mathbb{P}\{\omega\colon \xi_1(\omega) + \xi_2(\omega) = n\} = F_{\xi_1+\xi_2}(n+0) - F_{\xi_1+\xi_2}(n) =$$

$$= \sum\nolimits_{m=-\infty}^{+\infty} (F_1(n-m+0) - F_1(n-m))p_2(m) =$$

$$= \sum\nolimits_{m=-\infty}^{+\infty} p_1(n-m)p_2(m) \stackrel{def}{=} (p_1 * p_2)(n).$$

$p_1 * p_2$ — згортка дискретних розподілів.

Отже, довели таке твердження.

Твердження (згортка цілочисельних розподілів)

Нехай
ξ_1, ξ_2 — незалежні цілочисельні випадкові величини з розподілами p_1, p_2, відповідно. Тоді розподіл суми $\xi_1 + \xi_2$ є згорткою розподілів доданків

$$(\forall n \in \mathbb{Z}\colon\ )\ \mathbb{P}\{\omega\colon \xi_1(\omega) + \xi_2(\omega) = n\} = (p_1 * p_2)(n).$$

Твердження (теорема додавання розподілів Пуассона)

Нехай ξ_1, ξ_2 — незалежні випадкові величини, $\xi_j \in \mathcal{P}oiss(\lambda_j)$. Тоді

$$\xi = \xi_1 + \xi_2 \in \mathcal{P}oiss(\lambda_1 + \lambda_2).$$

Доведення. Оскільки $p_j(k) = \mathbb{P}\{\omega\colon \xi_j(\omega) = k\} = \dfrac{\lambda_j^k}{k!} e^{-\lambda_j}$, то за формулою згортки для цілочисельних дискретних розподілів

$$p(n) = \frac{1}{n!} \sum_{m=0}^{n} \frac{n!}{(n-m)!m!} \lambda_1^{n-m} \lambda_2^m \cdot e^{-(\lambda_1+\lambda_2)} = \frac{1}{n!} (\lambda_1 + \lambda_2)^n e^{-(\lambda_1+\lambda_2)}.$$