Київський національний університет імені Тараса Шевченка

Факультет радіофізики, електроніки і комп'ютерних систем



КОНСПЕКТ ЛЕКЦІЙ З КУРСУ «ТЕОРІЯ ЙМОВІРНОСТЕЙ»

Київ – 2019

УДК 519.22.35(075.8)
ББК 22.172я73

Рецензенти:

Єфіменко С.В.

Конспект лекцій з курсу «Теорія ймовірностей»

Методичний посібник призначений для студентів факультету радіофізики, електроніки і комп'ютерних систем для вивчення курсу «Теорія ймовірностей». Посібник повністю відповідає програмі курсу і містить тексти лекційних занять.

Затверджено вченою радою факультету радіофізики,
електроніки і комп'ютерних систем
(Протокол №4 від 12 листопада 2018 р.)

УДК 519.22.35(075.8)
ББК 22.172я73

© Єфіменко С.В., 2019
© Видавнича лабораторія факультету радіофізики, електроніки і комп'ютерних систем Київського університету імені Тараса Шевченка

<u>Лекція 1</u>. Предмет теорії ймовірностей. Випадкові події, їх ймовірності. Аксіоматика теорії ймовірностей. Елементи комбінаторики.

1. Що вивчає теорія ймовірностей?

Будь-яка наука базується на деяких початкових поняттях. Часто вони не визначаються, а вважаються інтуїтивно зрозумілими. В геометрії Евкліда до таких понять відносяться, наприклад, точка та пряма. З теорією ймовірностей ситуація складніша, тому що термінами, якими оперує ця наука, ми користуємось повсякденно. Ми говоримо про ймовірності можливих у майбутньому подій, отже, маємо на увазі певний фактор випадковості. І хоча всі слова для нас зрозумілі, проте спроба пояснити їх зміст може видатись проблематичною. Фактор випадковості очевидний, коли ми говоримо про карти, фішки, рух молекули, відмови приладів, аварії, катастрофи, примхи погоди. Важливо зрозуміти, відомості про які явища можна систематизувати та зробити їх предметом наукового вивчення.

Власне, теорія ймовірностей і починалась з азартних ігор, як широкого поля для спостережень за явищами з непередбачуваними наслідками. Виникнення та становлення цієї науки датують серединою 17-го сторіччя та пов’язують з іменами Блеза Паскаля, П’єра Ферма, Християна Гюйгенса та Якоба Бернуллі. Висновки, зроблені великими вченими, полягали у тому, що ***масові випадкові явища демонструють певні закономірності***. Так зрозуміло, що хоча й неможливо наперед точно передбачити, якою стороною впаде підкинута вгору монета, проте можна стверджувати, що при достатній кількості повторень цього експерименту ви побачите герб приблизно в половині випадків. І чим більш наполегливим виявиться експериментатор, тим точніше він одержить значення 0,5 для відносної кількості випадінь герба.

Отже, ***теорія ймовірностей вивчає закономірності випадкових явищ***. Проте не будь-яких явищ, а саме таких, які можуть бути багаторазово повторені при незмінних умовах і демонструють при цьому так звану ***статистичну стійкість***, тобто відносна кількість появи даного явища прямує до сталого значення із збільшенням кількості спостережень.

Сучасний розвиток теорії ймовірностей пов'язаний з багатьма галузями знань від квантової механіки та статистичної фізики до стратегій фінансових ринків, страхування, економіки, теорії хаосу, тощо. Неможливо не згадати про величезний внесок у розвиток теорії ймовірностей колишніх радянських вчених, зокрема і вчених київської школи (Б.В. Гнеденко, А.В.Скороход, М.Й. Ядренкло). Наука про випадкові явища ще з часів Пафнутія Львовича *Чебишова* (1821-1894) вважалась «російською» наукою. Її примножували багато вчених, проте сучасним своїм виглядом вона повністю завдячує Андрію Миколайовичу Колмогорову, який на початку 30-х років минулого сторіччя сформулював ***систему аксіом теорії ймовірностей***.

2. Випадкові події та їх ймовірності.

Відправним поняттям теорії ймовірностей є ***стохастичний експеримент*** – дія, яку можна повторювати багато разів, спостерігаючи за її випадковими наслідками. Прикладами найпростішого стохастичного експерименту можуть бути: підкидання монети або грального кубика, витягання карти з колоди, дані про народжуваність та смертність, влучання у мішень, тощо. Множину ***всіх можливих наслідків*** стохастичного експерименту називають ***простором елементарними***

подій. Цей простір позначають великою грецькою літерою омега: $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots\}$. Здається розумним кожній елементарній події ω_k певним чином поставити у відповідність число $p_k \in [0, 1]$, яке визначається стохастичним експериментом, так, щоб виконувалась рівність $\sum_k p_k = 1$. Число p_k і буде ймовірністю даної елементарної події ω_k. Правда, така схема діє лише у випадку, коли простір елементарних подій Ω ***дискретний***, тобто ***скінчений*** або ***зліченний***. Тоді для будь-якої складної події $A \subset \Omega$ (вона називається ***випадковою подією***) її ймовірність визначається так: $P(A) = \sum_{\omega_k \in A} p_k$. Множина ***всіх можливих випадкових подій*** позначатиметься $\Im$, а трійка об'єктів $(\Omega, \Im, P)$ називається ***імовірнісним простором***. У випадку, коли простір Ω континуальний, ситуація дещо складніша і буде з'ясована пізніше. Слід зазначити, що таким чином ми прискореним темпом проходимо шлях розвитку теоретико-імовірнісної науки від XVII до XIX сторічь. Першими досліджувались експерименти із ***скінченими*** множинами Ω, в які входять ***рівноможливі*** елементарні події. Такі експерименти описуються так званою ***<u>класичною імовірнісною моделлю</u>***. Її сенс полягає у тому, що якщо в експерименті відбуваються n <u>рівноможливих</u> наслідків, то кожному приписується ймовірність $1/n$, а ймовірність будь-якої випадкової події визначається через кількість k елементарних подій, що їй сприяють: $P(A) = |A|/|\Omega| = k/n$. Тут знаком $|A|$ позначається потужність події A, для скінченої множини елементарних подій – це просто кількість елементів в ній.

Приклади.

1. Нехай 1 раз підкидається монета. Тоді, очевидно, простір Ω складається лише з двох елементарних подій – випадіння герба та випадіння цифри: $\Omega = \{Г, Ц\}$, причому $P(Г) = P(Ц) = 1/2$.
2. Нехай 1 раз підкидається гральний кубик. Тоді $\Omega = \{1,2,3,4,5,6\}$, і кожна елементарна подія (випадіння грані з відповідною цифрою на ній) має ймовірність $1/6$. В цьому експерименті можна розглядати і складні випадкові події, наприклад, розглянемо подію A – випадіння парної цифри, тоді $A = \{2,4,6\}$, отже, $P(A) = 3/6 = 1/2$. Можна розглянути випадкову подію B – випала цифра більша за 4, тоді $B = \{5,6\}$, отже, $P(B) = 2/6 = 1/3$. Очевидно також, що подія – випала цифра 8 – є ***неможливою*** і має ймовірність 0. Неможлива подія позначається знаком $\varnothing$. Сама множина Ω називається ***достовірною*** подією – це подія, яка відбувається завжди, наприклад, для даного експерименту це випадіння цифри від 1 до 6.
3. Нехай монета підкидається до першого з'явлення герба. Тоді простір елементарних подій нескінченний: $\Omega = \{Г, ЦГ, ЦЦГ, ЦЦЦГ, \ldots\}$, проте зліченний. (Цей стохастичний експеримент не підпадає під схему класичної імовірнісної моделі!). Позначимо елементарні події цього експерименту $\omega_k = \underbrace{Ц \ldots Ц}_{k\text{-}1} Г$, $k \geq 1$. Тоді неважко зрозуміти, що достатньо покласти $P(\omega_k) = 1/2^k$, щоб умова $\sum_{k=1}^{\infty} p_k = 1$ була виконана.

Зауваження. Зазначимо, що ймовірності елементарних подій можна визначити і інакше. Наприклад, якщо маємо «неправильний» кубик, одна з граней якого більш важка і тому вона опиняється внизу частіше, випадінню кожної цифри можна приписати, скажімо, ймовірності $1/2$, $1/10$, $1/10$, $1/10$, $1/10$, $1/10$. Ймовірності випадкових подій A та B з прикладу 2 теж відповідним чином зміняться, проте математична модель залишиться вірною. Таким чином, для фіксованого стохастичного експерименту імовірнісний простір $(\Omega, \Im, P)$ може бути визначений ***неоднозначно***.

Простір елементарних подій може бути незліченним, мати потужність континуум. Наприклад, для експерименту, коли точка кидається на відрізок [0,1], простір Ω – це всі точки відрізку. Тоді говорити про ймовірності окремих елементарних подій – влучання у певну точку відрізку, немає сенсу, а ймовірність випадкової події, що точка влучила у певний інтервал, логічно вимірювати довжиною цього інтервалу. В інших стохастичних експериментах такою мірою може слугувати площа або об'єм відповідних геометричних множин. В таких задачах маємо справу з «геометричними» ймовірностями. Наступний приклад ілюструє подібну ситуацію.

Приклад.

4. Розглянемо «задачу про зустріч»: дві особи домовляються про зустріч на інтервалі часу T. Тоді простір Ω описується точками квадрату із стороною T: $\Omega = \{(x, y): 0 \le x \le T,\ 0 \le y \le T\}$. (Тут x та y відповідно означають час прибуття першої та другої осіб). Якщо додати умову, що кожна особа чекає лише протягом деякого часу $\tau < T$, то можна визначити випадкову подію A – зустріч відбулась, як підмножину Ω. Дійсно, $A = \{(x, y) \in \Omega : |x - y| \le \tau\}$.

   Тоді $P(A) = {площа\ A}\big/{площа\ \Omega} = \dfrac{T^2 - (T-\tau)^2}{T^2} = 1 - \left(1 - \dfrac{\tau}{T}\right)^2$.

Подібні задачі беруть початок від відомої задачі про голку Бюффона: на поверхню, розліновану паралельними прямими на відстані a одна від одної кидається голка довжиною $l < a$. З якою ймовірністю голка перетне одну з прямих? (Відповідь: $2l/\pi a$).

Слід зазначити, що геометричні ймовірності можуть привести до ще більшої неоднозначності при визначенні ймовірностей. Відомий ***парадокс Бертрана***, який стосується наступної задачі: яка ймовірність що навмання вибрана хорда круга більша за сторону вписаного у цей круг правильного трикутника. Бертран запропонував 3 розв'язки, які приводили до трьох різних відповідей: $1/4$, $1/3$, $1/2$.

3. Дії з випадковими подіями.

Випадкові події є підмножинами простору елементарних подій Ω, тому дії з ними подібні до дій в теорії множин. Важливо, проте, розібратись з їх імовірнісною інтерпретацією. У таблиці нижче наведені всі можливі дії з подіями. Тут A та B – випадкові події : $A, B \subset \Omega$.

Слід зауважити, що знаки операцій над подіями такі самі як і в теорії множин, що й не дивно, адже, як зазначалось вище, випадкова подія – це деяка множина елементарних подій. Що стосується назв дій з випадковими подіями, то допустиме вживання як суто імовірнісних термінів, так і теоретико-множинних (вони зазначені у дужках)

Операція	Назва операції	Зміст результату операції
$A \cup B$ (інколи $A + B$)	Сума (об'єднання) подій A та B	Подія, яка полягає у тому, що відбулась принаймні одна з подій A або B
$A \cap B$ (інколи $A \cdot B$)	Добуток (перетин) подій A та B	Подія, яка полягає у тому, що відбулись обидві події і A, і B
$A \setminus B$ (інколи $A - B$)	Різниця подій A та B	Подія, яка полягає у тому, що подія A відбулась, а подія B не відбулась
$\overline{A}$	Подія протилежна до A	Подія, яка полягає у тому, що подія A не відбулась
$A = B$	Рівносильні (рівні) події A та B	Події A та B відбуваються, або не відбуваються одночасно
$A \subset B$	Подія A сприяє події B або подія B є наслідком A	Якщо відбувається подія A, то відбувається подія B
$\varnothing$	Неможлива подія	Подія, яка не відбувається ніколи (не плутати з подією ймовірності 0!)
Ω	Достовірна подія	Подія, яка відбувається завжди (не плутати з подією ймовірності 1!)

Неважко довести деякі корисні властивості операцій над подіями, наприклад,

1) $A \setminus B = A \cdot \overline{B} = A \setminus (A \cdot B)$ 2) $\overline{A \cup B} = \overline{A} \cap \overline{B}; \quad \overline{A \cap B} = \overline{A} \cup \overline{B}$ (правила двоїстості)

3) $(A \cup B) \cdot C = (A \cdot C) \cup (B \cdot C)$

Повернемось до прикладу 2. Тут $A \cup B$ – це подія, що випала або парна цифра, або цифра більша за 4, тобто $A \cup B = \{2,4,5,6\}$; $A \cap B$ – це подія, що випала парна цифра, причому більша за 4, тобто $A \cap B = \{6\}$; $A \setminus B$ – це подія, що випала парна цифра, причому не більша за 4, тобто $A \setminus B = \{2,4\}$ і нарешті $\overline{A}$ – випала непарна цифра: $\overline{A} = \{1,3,5\}$.

Означення 1. Випадкові події A та B називаються ***несумісними***, якщо вони не відбуваються одночасно, тобто $A \cap B = \varnothing$.

4. Аксіоматика теорії ймовірностей.

Цей розділ змушено вносить ложку дьогтю в доволі просту попередню розмову. Справа в тому, що у випадку скінчених просторів Ω множиною випадкових подій $\mathfrak{I}$ є просто множина всіх підмножин простору Ω (потужність такої множини $2^{|\Omega|}$, тобто вона скінчена. При спробі ж узагальнити ситуацію на випадок довільної множини Ω, необхідно подбати про те, щоб операції над подіями не виводили нас за межі множини всіх подій $\mathfrak{I}$. А оскільки $\mathfrak{I}$ може містити нескінчену кількість елементів, необхідно забезпечити виконання нескінченої кількості додавань та добутків подій. Більше того, необхідно слушним чином визначати ймовірності таких подій. Лад у цих справах наводить ***аксіоматика Колмогорова***. Суть її у тому, що декларуються властивості класу

випадкових подій $\Im$, а ймовірність випадкової події визначається як особлива функція – імовірнісна міра $P(.)$.

Означення 2. Нехай Ω – довільна множина. Клас A підмножин Ω називається ***алгеброю*** множин, якщо

1) $\Omega \in \mathsf{A}$; 2) $\forall A \in \mathsf{A} \Rightarrow \overline{A} \in \mathsf{A}$; 3) $\forall A, B \in \mathsf{A} \Rightarrow A \cup B \in \mathsf{A}$

Таким чином алгебра множин є замкненою відносно виконання операцій переходу до протилежної множини та об'єднання множин.

Приклад. Мінімальна алгебра складається всього з двох множин: $\mathsf{A} = \{\Omega, \varnothing\}$. Нехай A – деяка підмножина Ω. Тоді $\mathsf{A} = \{\Omega, \varnothing, A, \overline{A}\}$ – також алгебра.

Означення 3. Нехай Ω – довільна множина. Клас $\Im$ підмножин Ω називається σ***-алгеброю*** множин, якщо

A1) $\Omega \in \Im$; A2) $\forall A \in \Im \Rightarrow \overline{A} \in \Im$ A3) $\forall A_1, A_2, \ldots \in \Im \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \Im$

(σ-властивість)

Означення 4. Нехай Ω – простір елементарних подій деякого експерименту, $\Im$ – σ***-алгебра*** випадкових подій. Будемо говорити, що визначена імовірнісна міра (або просто ймовірність) $P: \Im \to [0,1]$, якщо

P1) $P(\Omega) = 1$;

P2) $\forall A \in \Im \Rightarrow P(A) \geq 0$

P3) Для довільної послідовності $A_1, A_2, \ldots \in \Im$ попарно несумісних випадкових подій, тобто таких, що $A_i \cap A_j = \varnothing \quad \forall i \neq j$

$\Rightarrow P\left(\bigcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} P(A_k)$ (властивість σ-адитивності ймовірності)

Зауваження. Система аксіом A1-A3, P1-P3 не суперечна, проте неповна. Як уже зазначалось, є певна довільність у визначенні ймовірності P, отже, можливі різні імовірнісні $(\Omega, \Im, P)$ простори для одного й того ж самого стохастичного експерименту. Події із σ-алгебри $\Im$ називаються ***вимірними***.

Означення 5. Нехай $\Omega = [a,b]$ (стохастичним експериментом є вибір навмання точки на цьому відрізку). Через $\mathsf{B}([a,b])$ позначається ***борелівська*** σ***-алгебра*** множин на відрізку $[a,b]$, якщо вона є мінімальною σ-алгеброю, породженою напіввідкритими інтервалами виду $[\alpha, \beta) \subset [a,b]$. (Тобто, в неї включені всі такі напівінтервали, та всі множини, що є наслідками операцій з означення 3). Визначимо ймовірність попадання точки в інтервал $[\alpha, \beta)$: $P([\alpha, \beta)) = \dfrac{\beta - \alpha}{b - a}$. Тоді (Ω, B, P) – імовірнісний простір з геометричною ймовірністю на відрізку $[a,b]$.

5. Основні комбінаторні формули.

Оскільки суттєва частина задач в теорії ймовірностей спирається на класичне означення ймовірності, комбінаторика дає цілу низку корисних означень та ідей.

Основний принцип комбінаторики міститься в наступному твердженні. Із m елементів першої множини та n елементів другої множини можна утворити рівно $m \cdot n$ пар, складених із елементів, взятих по одному відповідно з першої та другої множин.

Доводиться це твердження складанням двовимірної таблиці, в якій в рядках містяться елементи першої множини, а в стовпчиках – елементи другої. Перехрестя

рядків та стовпчиків (їх рівно $m \cdot n$) визначать всі можливі пари. Цей принцип узагальнюється методом математичної індукції на довільну кількість множин елементів.

Наведемо основні комбінаторні схеми, які виникають в класичних задачах теорії ймовірностей.

1. ***Кількість перестановок*** із n елементів позначається P_n і рівна $P_n = n!$, оскільки на перше місце можна вибрати елемент n способами, на друге місце – n-1 способом, і так далі, аж поки на останнє місце не залишиться єдиний елемент.
2. ***Кількість розміщень*** із n елементів по k місцях, де $k \le n$ позначається A_n^k і рівна $A_n^k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \dfrac{n!}{(n-k)!}$ із тих самих міркувань. До цієї ж величини приводить підрахунок кількості виборів по k елементів із n можливих без повернення.
3. ***Кількість сполучень (комбінацій)*** із n елементів в групи по k елементів позначається C_n^k і рівна $C_n^k = \dfrac{n!}{k!(n-k)!}$, оскільки величину A_n^k треба зменшити у $k!$ разів, врахувавши несуттєвість порядку у групах по k елементів.
4. ***Кількість виборів по k елементів із n можливих з поверненням*** рівна, очевидно, n^k.
5. ***Кількість сполучень (комбінацій) з повтореннями*** із n елементів в групи по k елементів рівна C_{n+k-1}^k.

Приклади.

1. Нехай із колоди з 36 карт по черзі (тобто, враховуючи порядок карт) навмання витягаються 3 карти. Яка ймовірність, що всі вони пікової масті?
   При витяганні карт навмання ймовірності витягти будь-яку з них однакові, тобто маємо класичну схему. Різних способів витягти означеним чином 3 карти з колоди існує $36 \cdot 35 \cdot 34$. Нас цікавлять лише карти пікової масті, їх 9, отже, витягти 3 пікові карти ми можемо $9 \cdot 8 \cdot 7$ способами, тому ймовірність цієї події рівна $\dfrac{9 \cdot 8 \cdot 7}{36 \cdot 35 \cdot 34} \approx 0{,}012$.
2. Нехай так само, як і в попередньому випадку витягаємо карти, проте нехтуємо порядком їх з'явлення у трійці. Чи зміниться ймовірність витягти 3 пікові карти?
   Позначимо подію, яка полягає у витяганні трьох пікових карт через A. Для цієї задачі маємо $|\Omega| = C_{36}^3 = \dfrac{36 \cdot 35 \cdot 34}{1 \cdot 2 \cdot 3} = 7140$, $|A| = C_9^3 = \dfrac{9 \cdot 8 \cdot 7}{1 \cdot 2 \cdot 3} = 84$, $P(A) = \dfrac{84}{7140} \approx 0{,}012$. Як бачимо ймовірність не змінилась.
3. Із троянд червоного, білого та рожевого кольору складаються всі можливі букети по 3 квітки. Яка ймовірність, що навмання вибраний букет буде із червоних троянд.
   При складанні букетів кольори квітів можуть повторюватись. Отже, маємо сполучення із трьох кольорів у букети по 3 квітки з повтореннями. Їх кількість

рівна $C_{3+3-1}^{3-1} = C_5^2 = 10$. Тому ймовірність одержати букет із червоних троянд рівна 1/10, оскільки такий букет тільки один.

Лекція 2. Властивості ймовірності. Умовні ймовірності, незалежні події. Формула повної ймовірності та формули Байеса.

1. Властивості ймовірності, що випливають з аксіом

В попередній лекції було сформульовано систему аксіом А1-А3, Р1-Р3, яка формалізує поняття випадкової події та її ймовірності. Зараз вивчимо властивості імовірнісної міри, які безпосередньо випливають із аксіом, що її визначають.

Нехай $(\Omega, \Im, P)$ – імовірнісний простір, де Ω – простір елементарних подій даного експерименту, $\Im$ – σ-алгебра випадкових (вимірних) подій, P – ймовірність випадкових подій із $\Im$.

Теорема 1. Для будь-якої випадкової події $A \in \Im$: $P(\overline{A}) = 1 - P(A)$. (1)

Доведення. Безпосередньо з аксіоми Р3 випливає, що $P(\overline{A}) + P(A) = P(\Omega) = 1$ оскільки події A та $\overline{A}$ несумісні.

Теорема 2. $P(\varnothing) = 0$.

Доведення. Дійсно, з попереднього маємо: $P(\varnothing) = 1 - P(\Omega) = 0$. (2)

Теорема 3. Якщо подія B є наслідком події A: $A \subseteq B$, то для ймовірності різниці цих подій справедлива рівність $P(B \setminus A) = P(B) - P(A)$. (3)

Доведення. Справді, для таких подій вірно, що $B = A \cup (B \setminus A)$, причому події A та $B \setminus A$ несумісні, отже, $P(B) = P(A) + P(B \setminus A)$, звідки й випливає твердження.

Теорема 4. Для довільних подій $A, B \in \Im$:

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

Доведення. Подію $A \cup B$ легко подати у вигляді суми попарно несумісних подій: $A \cap B = (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B)$, причому як подія A, так і подія B є наслідками події $A \cap B$. Тому з аксіоми Р3 та теореми 3 маємо: $P(A \cup B) = P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B) = P(A) + P(B) - P(A \cap B)$.

Зауваження. Теорема 4 може бути узагальнена на довільну кількість n випадкових подій $A_1, A_2, \ldots, A_n \in \Im$:

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k) - \sum_{\substack{k_1, k_2 = 1 \\ (k_1 \neq k_2)}}^{n} P(A_{k_1} \cap A_{k_2}) + \cdots + (-1)^{n-1} P\left(\bigcap_{k=1}^{n} A_k\right). \quad (5)$$

Зокрема, для трьох випадкових подій маємо формулу:

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

Доведення теореми 4 (її ще називають ***теоремою додавання***), наприклад, методом математичної індукції пропонується провести самостійно.

Приклад. Розглянемо так звану задачу «про збіг», а саме: n чисел випадковим чином переставляють на n місцях. Яка ймовірність, що принаймні одне з чисел $1, 2, \ldots, n$ виявиться на своєму місці? (Ця задача відома і з більш ліричним змістом: дехто випадковим чином розкладає n листів у n конвертів. Знайти ймовірність, що принаймні один адресат одержить «свій» лист.)

Позначимо A_k – подію, що конкретне число k опиниться на місці з номером

k. Тоді $A=\bigcup_{k=1}^{n}A_k$ – подія, що принаймні одне з чисел $1,2,\ldots,n$ виявиться на своєму місці. Вважаючи, що всі перестановки чисел $1,2,\ldots,n$ рівно можливі (тобто маємо класичну ймовірнісну модель), легко одержати, що $P(A_k)=\dfrac{(n-1)!}{n!}=\dfrac{1}{n}$ для будь-якої події A_k. Аналогічно, $P(A_{k_1}\cap A_{k_2})=\dfrac{(n-2)!}{n!}=\dfrac{1}{n(n-1)}$ для довільної пари подій A_{k_1}, A_{k_2}. Продовжуючи, одержимо, що з формули (5) безпосередньо випливає:

$$P(A)=\sum_{k=1}^{n}\frac{1}{n}-\sum_{\substack{k_1,k_2=1\\(k_1\neq k_2)}}^{n}\frac{1}{n(n-1)}+\cdots+(-1)^{n-1}\frac{1}{n!}=1-C_n^2\frac{(n-2)!}{n!}+\cdots+(-1)^{n-1}\frac{1}{n!}=$$

$$=1-\frac{1}{2!}+\frac{1}{3!}-\cdots+(-1)^{n-1}\frac{1}{n!}$$

А прийнявши до уваги представлення функції e^x при $x=-1$ у вигляді ряду Тейлора, бачимо, що при $n\to\infty$: $P(A)\to 1-\dfrac{1}{e}\approx 0{,}632$

Теорема 5 (властивість неперервності ймовірності). Для монотонно зростаючої послідовності випадкових подій $A_1\subset A_2\subset\ldots\subset A_n\subset\ldots\in\mathfrak{I}$:

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}P(A_n). \qquad (6)$$

Для монотонно спадної послідовності випадкових подій $A_1\supset A_2\supset\ldots\supset A_n\supset\ldots\in\mathfrak{I}$: $$P\left(\bigcap_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}P(A_n). \qquad (7)$$

Доведення. Доведемо першу частину твердження. Розглянемо події $B_1=A_1, B_2=A_2\setminus A_1,\ldots,B_n=A_n\setminus A_{n-1},\ldots$ Оскільки вони попарно несумісні, то згідно аксіомі Р3, маємо: $P\left(\bigcup_{n=1}^{\infty}B_n\right)=\sum_{n=1}^{\infty}P(B_k)=\lim_{n\to\infty}\sum_{k=1}^{n}P(B_k)=$

$=\lim_{n\to\infty}\left(P(A_1)+\sum_{k=1}^{n-1}(P(A_{k+1})-P(A_k))\right)=\lim_{n\to\infty}P(A_n)$. Аналогічно доводиться і друга частина твердження.

Зауваження. Якщо ввести позначення $\bigcup_{n=1}^{\infty}A_n=\lim_{n\to\infty}A_n$ для суми монотонно зростаючої послідовності подій та $\bigcap_{n=1}^{\infty}A_n=\lim_{n\to\infty}A_n$ відповідно для добутку монотонно спадної послідовності випадкових подій, то властивості (7) та (8) можна об'єднати в одну формулу, яка більш наочно передає зміст властивості неперервності ймовірності, а саме: $$P\left(\lim_{n\to\infty}A_n\right)=\lim_{n\to\infty}P(A_n)$$

2. Умовні ймовірності, незалежні події. Формула повної ймовірності та формули Байеса.

Нехай в деякому стохастичному експерименті відбулася випадкова подія B. Як впливатиме цей факт на ймовірності інших подій? Спробуємо з'ясувати це спочатку на прикладі.

Нехай відомо, що при двох підкиданнях кубика на першому з них випала цифра «1». Позначимо цю подію через B і з'ясуємо, якою буде за цієї умови ймовірність події A – сума цифр парна. Таку умовну ймовірність події A при умові, що відбулась подія B позначають як $P(A/B)$. Зрозуміло, що $B=\{"11","12","13","14","15","16"\}$ і $|B|=6$. За умови, що відбулась подія B, події A сприяють лише 3 елементарні події – $\{"11","13","15"\}=A\cap B$, тому $P(A/B)=\dfrac{|A\cap B|}{|B|}=\dfrac{3}{6}=\dfrac{1}{2}$. З іншого боку, ця відповідь може бути визначена, як

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

Означення 1. Нехай $(\Omega,\Im,P)$ – імовірнісний простір, пов'язаний з деяким стохастичним експериментом. Умовна ймовірність події A при умові, що сталась подія B (тобто $P(B)\neq 0$) визначається формулою:

$$P(A/B)=\frac{P(A\cap B)}{P(B)} \qquad (8)$$

Зауваження. З очевидних властивостей умовної ймовірності, а саме:

1) $P(A/B)\geq 0 \quad \forall A\in\Im$; 2) $P(\Omega/B)=1$; 3) $P\left(\bigcup\limits_{n=1}^{\infty}A_n/B\right)=\sum\limits_{n=1}^{\infty}P(A_n/B)$ для довільних попарно несумісних подій $A_n\in\Im$;

випливає, що при фіксованій події B $(\Omega,\Im,P(\cdot/B))$ – також імовірнісний простір.

З формули (8) одержимо, що $P(A\cap B)=P(B)\cdot P(A/B)$ (9)

Ця формула (9) носить назву ***формули множення***.

Теорема 1. Формула множення може бути узагальнена для довільної кількості n випадкових подій $A_1,A_2,\ldots,A_n\in\Im$:

$$P\left(\bigcap_{k=1}^{n}A_k\right)=P(A_1)\cdot P(A_2/A_1)\cdot P(A_3/A_1\cap A_2)\cdot\ldots\cdot P(A_n/(A_1\cap A_2\cap\cdots\cap A_{n-1})) \qquad (10)$$

Доведення цієї теореми пропонується провести самостійно методом математичної індукції.

Цікавим є питання, які події не впливають одна на одну, тобто відбуваються або не відбуваються незалежно? Всі інтуїтивно розуміють, що означає незалежність подій – наприклад, результат окремого підкидання кубика чи монети не повинен впливати на наступні результати. Тобто розглядаючи дві події A та B, природно вважати їх незалежними, якщо $P(A/B)=P(A)$ або $P(B/A)=P(B)$. З означення умовної ймовірності випливає тоді, що $P(A\cap B)=P(A)P(B)$.

Означення 2. Випадкові події A та B називаються ***стохастично незалежними*** (або просто ***незалежними***), якщо $P(A\cap B)=P(A)P(B)$ (11)

Зауваження. За умови, що $P(A)\neq 0, P(B)\neq 0$, $P(A\cap B)=P(A)P(B)\Leftrightarrow P(A/B)=P(A)\Leftrightarrow P(B/A)=P(B)$.

Означення 3. Випадкові події $A_1, A_2,\ldots, A_n\in\Im$ називаються ***незалежними у сукупності***, якщо $\forall k\in\{1,2,\ldots,n\}$ і $\forall A_{i_1},\ldots,A_{i_k}$, де всі індекси належать множині $\{1,2,\ldots,n\}$, має місце рівність:

$$P(A_{i_1}\cap\ldots\cap A_{i_k})=P(A_{i_1})\cdot\ldots\cdot P(A_{i_k}) \qquad (12)$$

Зауваження. Якщо $k=2$, рівність (12) виражає попарну незалежність. Зрозуміло, що попарна незалежність недостатня для незалежності у сукупності.

Зауваження. Не плутайте співзвучні поняття ***незалежності*** та ***несумісності*** випадкових подій! Більш того, якщо маємо несумісні випадкові подій A та B $P(A)>0, P(B)>0$, то A та B – залежні: $P(A\cap B)=0\neq P(A)P(B)>0$

Приклад 1. (приклад С.Н.Бернштейна демонструє різницю між попарною незалежністю та незалежністю у сукупності). Розглянемо правильний тетраедр, три грані якого пофарбовані відповідно у червоний, зелений та синій кольори, на четверту грань нанесені всі три вказані кольори. Припустимою що тетраедр кидається на площину. Нехай випадкові події A, B та C – означають що площини торкнулась грань, на якій присутні відповідно червоний, зелений та синій кольори. Тоді $P(A)=P(B)=P(C)=\frac{2}{4}=\frac{1}{2}$. Далі, $P(A\cap B)=P(B\cap C)=P(C\cap A)=\frac{1}{4}=\frac{1}{2}\cdot\frac{1}{2}$, що означає попарну незалежність подій A, B та C. Проте $P(A\cap B\cap C)=\frac{1}{4}\neq\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=P(A)\cdot P(B)\cdot P(C)$, тобто події A, B та C не є незалежними в сукупності.

Означення 4. Випадкові події H_1, $H_2,\ldots$, $H_n\in\Im$ утворюють ***повну групу подій***, якщо:

1) $\bigcup_{k=1}^{n}H_k=\Omega$; 2) $P(H_k)>0\ \forall k=\overline{1,n}$; 3) $H_i\cap H_j=\varnothing\ \forall i\neq j$

Теорема 2 (Формула повної ймовірності). Для повної групи подій $H_1, H_2,\ldots, H_n\in\Im$ та ймовірності довільної випадкової події $A\in\Im$ має місце формула:

$$P(A)=\sum_{k=1}^{n}P(A/H_k)\cdot P(H_k) \qquad (13)$$

Доведення. $P(A)=P(A\cap\Omega)=P\left(A\cap\left(\bigcup_{k=1}^{n}H_k\right)\right)=P\left(\bigcup_{k=1}^{n}(A\cap H_k)\right)=\sum_{k=1}^{n}P(A\cap H_k)$.

Далі скористаємось формулою множення (9): $P(A\cap H_k)=P(H_k)\cdot P(A/H_k)$ і одержимо твердження теореми.

Формула повної ймовірності (13) дозволяє ***безумовну*** ймовірність деякої події A визначити через ***умовні*** відносно набору деяких гіпотез ймовірності, які буває як раз просто визначити в умовах задачі.

Приклад 2. Нехай група студентів складає залік, для одержання якого досить правильної відповіді на одне питання випадково витягнутого білета. Припустимо, що з N питань є n $(n<N)$ «нещасливих», відповіді на які не знає жоден студент. Відповіді на решту питань знають всі студенти. Знайти ймовірність здати залік для першого та другого студента (вважаємо, що вони послідовно витягають білети).

Позначимо через H_1 випадкову подію, яка полягає в тому, що перший студент склав залік (вона означає, що студент витяг «щасливий» білет) і через $H_2 = \overline{H_1}$ – що не склав (тобто витяг «нещасливий» білет). Очевидно, що $P(H_1) = \frac{N-n}{N}$, $P(H_2) = \frac{n}{N}$. Знайдемо ймовірність випадкової події A – що залік одержав другий студент. Очевидно, що ймовірність цієї події неможливо визначити безпосередньо, адже вона відбувається або не відбувається після того, як сталася одна з подій H_1 або H_2. Скористаємось формулою повної ймовірності, адже події H_1 та H_2 утворюють повну групу подій, а умовні ймовірності $P(A/H_k), k = 1,2$ легко визначити. Отже,

$$P(A) = P(A/H_1)\cdot P(H_1) + P(A/H_2)\cdot P(H_2) = \frac{N-n-1}{N-1}\cdot\frac{N-n}{N} + \frac{N-n}{N-1}\cdot\frac{n}{N} =$$
$$= \frac{N-n}{(N-1)N}(N-n-1+n) = \frac{N-n}{N}$$

На перший погляд видається дивним, що ймовірності витягнути «щасливий» білет, а отже, здати залік однакові і для першого, і для другого студента. Проте, це лише стохастичне підтвердження загальної справедливості. Переконайтесь, що взагалі, ймовірність скласти залік однакова для кожного студента, незалежно від того, яким по черзі він витягає білет.

Приклад 3. Двоє стрільців можуть влучити у мішень відповідно із ймовірностями 0,1 та 0,8. Вони підкидають монету, щоб визначити, хто буде стріляти, і один з них робить постріл по мішені. Яка ймовірність влучити в мішень при даному пострілі.

Оскільки невідомо, який стрілець робить постріл, зробимо всі можливі припущення щодо цього та використаємо формулу повної ймовірності. Отже, нехай H_1 та H_2 – випадкові події, що постріл зроблений відповідно першим або другим стрільцем. Вони утворюють повну групу подій, і крім того, $P(H_1) = P(H_2) = \frac{1}{2}$. Нехай випадкова подія A – це влучання в мішень. Тоді $P(A/H_1) = 0{,}1$ і $P(A/H_2) = 0{,}8$.

$$P(A) = P(A/H_1)\cdot P(H_1) + P(A/H_2)\cdot P(H_2) = 0{,}1\cdot 0{,}5 + 0{,}8\cdot 0{,}5 = 0{,}45.$$

Тепер з'ясуємо, як змінюються ймовірності подій, що входять до повної групи, якщо напевно відомо, що сталась деяка випадкова подія A. При цьому початкові ймовірності гіпотез $P(H_k)$ прийнято називати «апріорними ймовірностями» на відміну від «апостеріорних ймовірностей» $P(H_k/A)$.

Теорема 3 (Формули Байеса). Для апостеріорних ймовірностей повної групи подій $H_1, H_2, \ldots, H_n \in \Im$ та довільної випадкової події $A \in \Im$ $(P(A) \neq 0)$ справедливі формули:

$$P(H_k/A) = \frac{P(A/H_k)\cdot P(H_k)}{\sum_{i=1}^{n} P(A/H_i)\cdot P(H_i)} \quad \forall k = \overline{1,n} \qquad (14)$$

Доведення. З означення умовної ймовірності для довільної гіпотези $H_k, k = \overline{1,n}$ маємо: $P(H_k/A) = \frac{P(H_k \cap A)}{P(A)}$. Далі використаємо формули множення (11) та

повної ймовірності (13): $P(H_k / A) = \dfrac{P(A/H_k)\cdot P(H_k)}{P(A)} = \dfrac{P(A/H_k)\cdot P(H_k)}{\sum_{i=1}^{n} P(A/H_i)\cdot P(H_i)}$.

Приклад 4. В умовах прикладу 3 визначити ймовірності випадкових подій, що постріл зроблений відповідно першим та другим стрільцем, якщо відомо, що при пострілі стався промах.

Нам необхідно визначити апостеріорні ймовірності $P(H_1/\overline{A})$ та $P(H_2/\overline{A})$. Оскільки $P(\overline{A}) = 1 - P(A) = 0{,}55$, то формули Байеса дають такі результати:

$$P(H_1/\overline{A}) = \frac{P(\overline{A}/H_1)\cdot P(H_1)}{P(\overline{A})} = \frac{0{,}9\cdot 0{,}5}{0{,}55} = \frac{9}{11};$$

$$P(H_2/\overline{A}) = \frac{P(\overline{A}/H_2)\cdot P(H_2)}{P(\overline{A})} = \frac{0{,}2\cdot 0{,}5}{0{,}55} = \frac{2}{11}.$$

Зверніть увагу, як змінились апріорні ймовірності, які , були рівні 0,5, якщо стало відомо, що стався промах.

Приклад 5. В урні знаходиться n куль білого та чорного кольорів. Всі припущення про склад урни вважати рівноможливими. Навмання витягнута куля виявилась білою. Обчислити ймовірності всіх гіпотез про склад урни. Яка з них найбільш ймовірна?

Позначимо через $H_0, H_1, \ldots, H_n$ випадкові гіпотези, де гіпотеза $H_k \ (k = \overline{0,n})$ означає, що в урні знаходиться рівно k білих куль. За умовою задачі $P(H_k) = \frac{1}{n+1}$. Сталась випадкова подія A – навмання витягнута куля виявилась білою. Тоді $P(A) = \sum_{k=1}^{n} P(A/H_k)\cdot P(H_k) = \sum_{k=1}^{n} \frac{k}{n}\cdot\frac{1}{n+1} = \frac{1}{n(n+1)}\sum_{k=1}^{n} k = \frac{n(n+1)}{2n(n+1)} = \frac{1}{2}$. Необхідно визначити ймовірності $P(H_k/A) \ \ \forall k = \overline{0,n}$.

Маємо: $P(H_k/A) = \dfrac{P(A/H_k)\cdot P(H_k)}{P(A)} = \frac{2k}{n(n+1)}$. Зрозуміло, що найбільш ймовірною при умові A є гіпотеза H_n (тобто, що всі кулі в урні – білі): $P(H_n/A) = \frac{2}{(n+1)}$.

Приклад 5. Яку найменшу кількість пострілів треба зробити стрільцеві, який здатен влучити в мішень у 7 випадках з 10, щоб ймовірність враження мішені була не менше 0,9?

Вважаючи постріли незалежними, знайдемо ймовірність події A – враження мішені стрільцем при n пострілах. Оскільки для цього досить принаймні одного влучання, то $P(\overline{A}) = (1-0{,}7)^n = (0{,}3)^n$, а отже, для відшукання значення n потрібно розв'язати нерівність $P(A) \ge 0{,}9$, тобто $1-(0{,}3)^n \ge 0{,}9$, звідки $(0{,}3)^n \le 0{,}1$. Тоді $n(\lg 3 - 1) \le -1$, або $n \ge \dfrac{1}{1-\lg 3}$. Найменше ціле , яке задовольняє цю нерівність – це число 2. Отже, двох пострілів досить, щоб вразити мішень з ймовірністю не меншою від 0,9.

Лекція 3.Послідовності незалежних в*и*пробувань Бернуллі. Граничні теореми.

1. Схема Бернуллі.

Розглянемо стохастичний експеримент, який полягає в багатократному повторенні деякого досліду. Вважатимемо, що множина можливих наслідків кожного досліду скінчена, а самі досліди відбуваються незалежно (в сукупності) і при однакових умовах. Такий стохастичний експеримент називають ***послідовністю незалежних однорідних випробувань***. Для спрощення міркувань розглянемо спочатку ситуацію, коли кожний дослід має лише два можливих варіанти завершення, які умовно позначимо, як «успіх» та «невдача», що стаються відповідно з ймовірностями p та q $(p+q=1)$. В такій постановці задачі цікавляться не результатами кожного окремого досліду, а кількістю «успіхів» в серії із n випробувань, приміром кількість влучань в серії пострілів. Саме таку задачу успішно дослідив Я. Бернуллі. На його честь її називають ***схемою Бернуллі***.

Теорема 1. В серії із n незалежних і однорідних випробувань ймовірність $P_n(k)$ події, що кількість «успіхів» рівна k $(0 \le k \le n)$, визначається формулою Бернуллі:

$$P_n(k) = C_n^k p^k q^{n-k} \qquad (1)$$

Набір ймовірностей $P_n(k)$ для $0 \le k \le n$ називається ***біноміальним*** (інколи кажуть ***біномним***) ***розподілом з параметрами*** n ***та*** p.

Доведення. Простір елементарних подій послідовності незалежних випробувань можна описати, як набір ланцюжків довжини n, складених із літер «У» та «Н», поставлених у відповідності з результатом кожного стохастичного експерименту. Зверніть увагу, що елементарні події даного стохастичного експерименту не є рівноможливими. Наприклад, при $n=2$ $\Omega = \{НН, УУ, НУ, УН\}$ і, взагалі кажучи, $P(НН) = q^2 \ne P(УУ) = p^2$. Знайдемо ймовірність події $A_n(i_1, i_2, \ldots, i_k)$, що успіхи мали місце лише у дослідах з номерами $i_1, i_2, \ldots, i_k$. Із-за незалежності дослідів маємо $P(A_n(i_1, i_2, \ldots, i_k)) = p^k q^{n-k}$, і дана ймовірність не залежить від номерів випробувань, у яких стались «успіхи», а визначається лише їх кількістю. Отже, ймовірність $P_n(k)$ є сумою ймовірностей $P(A_n(i_1, i_2, \ldots, i_k))$ по всіх можливих виборах номерів $i_1, i_2, \ldots, i_k$. Кількість таких виборів, очевидно, рівна C_n^k, що й доводить справедливість формули (1).

Зауваження 1. Назва даного розподілу пов'язана з тим фактом, що кожна ймовірність $P_n(k)$ є елементом біному: $(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = \sum_{k=0}^{n} P_n(k) = 1.$

Зауваження 2. Ця теорема може бути узагальнена на випадок довільної скінченої кількості варіантів завершення кожного випробування, а саме, має місце наступне твердження (доведення пропонується провести самостійно).

Теорема 2. Нехай в серії n незалежних однорідних випробувань кожне з них може завершитись випадковими наслідками $\nu_1, \nu_2, \ldots, \nu_m$ з ймовірностями, рівними відповідно $p_1, p_2, \ldots, p_m$ (при цьому $p_1 + p_2 + \ldots + p_m = 1$). Тоді ймовірність $P_n(k_1, k_2, \ldots, k_m)$ події, що наслідки $\nu_1, \nu_2, \ldots, \nu_m$ сталися відповідно $k_1, k_2, \ldots, k_m$

разів $(k_1 + k_2 + \ldots + k_m = n)$, є доданком розкладу полінома $(p_1 + p_2 + \ldots + p_m)^n$ і визначається формулою

$$P_n(k_1, k_2, \ldots, k_m) = \frac{n!}{k_1! k_2! \cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} \qquad (2)$$

Набір ймовірностей $P_n(k_1, k_2, \ldots, k_m)$ для всіх допустимих значень $k_1, k_2, \ldots, k_m$ називається ***поліноміальним розподілом***. Для них справедливо, що $\sum_{k_1, \ldots k_m} P_n(k_1, k_2, \ldots, k_m) = (p_1 + \ldots + p_m)^n = 1.$

Теорема 3. Ймовірність, що перший успіх в схемі Бернуллі станеться у в***и***пробуванні з номером k рівна $P(\tau = k) = pq^{k-1}$. (Тут τ означає випадковий номер першого успішного випробування в серії).

Доведення цього факту цілком очевидне – до першого успіху у випробуванні з номером k сталось рівно $k-1$ невдач, отже, $P(\tau = k) = pq^{k-1}$.

Набір ймовірностей pq^{k-1} для $0 \le k \le n$ називається ***геометричним розподілом з параметром*** p. Цей розподіл має важливу властивість – так звану «відсутність післядії»: $P(\tau = k + m / \tau \ge m) = P(\tau = k) = pq^{k-1}$. (Про цю властивість буде згадано і в наступній лекції).

Приклад. На деякому виробництві ймовірність виготовлення бракованого виробу складає $p = 0{,}01$. Визначити ймовірність, що в партії із 100 виробів бракованих буде рівно 2, не більше ніж 2. Яку партію виробів треба взяти, щоб ймовірність знаходження в ній принаймні одного бракованого виробу була більша за 0,5?

Ймовірність, що в партії із 100 виробів бракованих буде рівно 2, визначимо за формулою (1): $P_{100}(2) = C_{100}^2 p^2 q^{100-2} = \frac{100 \cdot 99}{2} (0{,}01)^2 (0{,}99)^{98} \approx 0{,}185$. Ймовірність, що в даній партії кількість бракованих не перевищує двох, знайдемо як суму ймовірностей $P_{100}(2) + P_{100}(1) + P_{100}(0) \approx 0{,}185 + 0{,}370 + 0{,}366 = 0{,}921$

Для відповіді на останнє запитання необхідно розв'язати відносно n нерівність $\sum_{k=1}^{n} P_n(k) = 1 - P_n(0) > 0{,}5$, або $P_n(0) < 0{,}5$, що означає виконання нерівності $(0{,}99)^n < 0{,}5$. Логарифмуючи, одержимо: $n > -\frac{\lg 2}{\lg 0{,}99}$, звідки $n \ge 69$.

Дослідимо тепер поведінку послідовності значень ймовірностей $P_n(k)$ біноміального розподілу при зміні k від 0 до n. З цією метою розглянемо відношення

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

Звідси видно, що при $(n+1)p < k$ виконується нерівність $P_n(k) < P_n(k-1)$, а при $(n+1)p > k$ – нерівність $P_n(k) > P_n(k-1)$. Таким чином, значення $P_n(k)$ спочатку зростають, а потім спадають, при цьому $\max_{0 \le k \le n} P_n(k) = P_n(k_0)$ досягається при значенні k рівному $k_0 = [(n+1)p]$ – ***найбільш ймовірна кількість «успіхів»,*** – і

називається ***максимальною ймовірністю***. Якщо ж число $(n+1)p$ – ціле, то максимальних значень ймовірності два: $P_n(k_0) = P_n(k_0 - 1).0$

Приклад. Яку найбільш ймовірну кількість м'ячів закине в кільце баскетболіст в серії із 50 кидків, якщо він влучає з ймовірністю 0,8 при кожному підкиданні м'яча, і яка ймовірність такої кількості закинутих м'ячів?

Маємо: $n = 50,\ p = 0{,}8$. Знайдемо $k_0 = [(n+1)p] = [51 \cdot 0{,}8] = 40$ Відповідна максимальна ймовірність дорівнює $P_{50}(40) = C_{50}^{40}\, 0{,}8^{40} 0{,}2^{10} \approx 0{,}14$.

2. Граничні теореми в серії незалежних випробувань Бернуллі.

При великих кількостях незалежних випробувань безпосередні обчислення за формулою (1) стають обтяжливими, і виникає питання про використання наближених формул. Перші граничні теореми для незалежних випробувань Бернуллі довів А. Муавр у 1733 р. в припущенні, що $p = q = 0{,}5$.

Введемо наступні позначення: $\Phi(x) = \int\limits_{-\infty}^{x} \varphi(t)dt$ – функція нормального розподілу, $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ – щільність нормального розподілу. Властивості та зміст цих важливих функцій будуть визначені дещо пізніше.

Теорема 4 (Гранична теорема Муавра-Лапласа). Припустимо, що в схемі Бернуллі p та q означають ймовірності «успіху» та «невдачі» відповідно, нехай $n \to \infty$, причому k та $n-k$ також досить великі. Тоді

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{k-np}{\sqrt{npq}}\right) \quad \text{– локальна теорема Мавра-Лапласа} \tag{4}$$

$$P(n_1 \le k \le n_2) = \sum_{k=n_1}^{n_2} P_n(k) \approx \Phi\left(\frac{n_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{n_1 - np}{\sqrt{npq}}\right) \tag{5}$$

– інтегральна теорема Мавра-Лапласа.

Доведення. Для доведення скористаємося формулою Стірлінга: $n! \approx \sqrt{2\pi n} n^n e^{-n}$, справедливою при великих n.

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k} \approx \frac{\sqrt{2\pi n} n^n e^{-n} p^k q^{n-k}}{\sqrt{2\pi k} k^k e^{-k} \sqrt{2\pi(n-k)} (n-k)^{n-k} e^{-n+k}} =$$

$$= \frac{\sqrt{n} n^n p^k q^{n-k}}{\sqrt{2\pi k(n-k)} k^k (n-k)^{n-k}} = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{k}{np}\right)^{-k} \left(\frac{n-k}{nq}\right)^{-(n-k)}$$

Позначимо через $x_k \overset{\text{def}}{=} x = \frac{k-np}{\sqrt{npq}}$. Якщо значення x знаходиться у скінченому проміжку, то k та $n-k$ зростають разом з n. Тоді $k = np + x\sqrt{npq}$, а $n - k = n - np - x\sqrt{npq} = nq - x\sqrt{npq}$. Враховуючи це, розглянемо логарифм виразу $R_n = \left(\frac{k}{np}\right)^{-k} \left(\frac{n-k}{nq}\right)^{-(n-k)}$: $\ln R_n = -k \ln\left(\frac{k}{np}\right) - (n-k)\ln\left(\frac{n-k}{nq}\right) =$

$$= -\left(np + x\sqrt{npq}\right)\ln\left(1 + x\sqrt{\frac{q}{np}}\right) - \left(nq - x\sqrt{npq}\right)\ln\left(1 - x\sqrt{\frac{p}{nq}}\right)$$

Далі скористаємось формулою Тейлора для функції $\ln(1+t) \approx t - \frac{t^2}{2} + O\left(t^3\right)$ при $t \to 0$:

$$\ln R_n \approx -\left(np + x\sqrt{npq}\right)\left(x\sqrt{\frac{q}{np}} - \frac{x^2 q}{2np} + O\left(\frac{1}{n^{3/2}}\right)\right) - \left(nq - x\sqrt{npq}\right)\left(-x\sqrt{\frac{p}{nq}} - \frac{x^2 p}{2nq} + O\left(\frac{1}{n^{3/2}}\right)\right) =$$

$$= -x\sqrt{npq} - x^2 q + \frac{x^2 q}{2} + O\left(\frac{1}{\sqrt{n}}\right) + x\sqrt{npq} - x^2 p + \frac{x^2 p}{2} + O\left(\frac{1}{\sqrt{n}}\right) = -\frac{x^2}{2} + O\left(\frac{1}{\sqrt{n}}\right)$$

Таким чином, $R_n \approx e^{-\frac{x^2}{2}}, n \to \infty$.

Крім того, $\sqrt{\frac{n}{k(n-k)}} = \left(\left(\frac{k}{n}\right)\left(1 - \frac{k}{n}\right)n\right)^{-1/2} \approx \left(np(1-p)\right)^{-1/2} = \frac{1}{\sqrt{npq}}, n \to \infty$. Остаточно одержимо, що при $n \to \infty$ $P_n(k) \approx \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{npq}}\varphi(x) = \frac{1}{\sqrt{npq}}\varphi\left(\frac{k-np}{\sqrt{npq}}\right)$.

Тим самим твердження локальної граничної теореми доведено. (Правда, ми тут опускаємо дослідження оцінка точності збіжності).

Покажемо, звідки випливає справедливість інтегральної граничної теореми Муавра-Лапласа. Розглянемо ймовірність:

$$P(n_1 \le k \le n_2) = \sum_{k=n_1}^{n_2} P_n(k) \approx \sum_{k=n_1}^{n_2} \frac{1}{\sqrt{npq}}\varphi\left(\frac{k-np}{\sqrt{npq}}\right) = \sum_{k=n_1}^{n_2} \frac{1}{\sqrt{npq}}\varphi(x_k)$$

Оскільки $x_{k+1} - x_k = \frac{1}{\sqrt{npq}}$, то при $n \to \infty$ ймовірність $P(n_1 \le k \le n_2)$ є інтегральною сумою, що прямуватиме до інтегралу $\int_a^b \varphi(x)dx$, де $a = \frac{n_1 - np}{\sqrt{npq}}$, $b = \frac{n_2 - np}{\sqrt{npq}}$. Враховуючи прийняті позначення, одержимо, що

$$P(n_1 \le k \le n_2) \approx \Phi\left(\frac{n_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{n_1 - np}{\sqrt{npq}}\right).$$

Слід зазначити, що значення функцій $\Phi(x)$ та $\varphi(x)$, що дають граничні наближення, табульовані. При використанні таблицями варто враховувати властивості цих функцій (перевірте їх самостійно), а саме: $\varphi(-x) = \varphi(x)$, $\int_{-\infty}^{\infty} \varphi(x)dx = 1$, $\Phi(x) = 1/2 + \operatorname{sgn}(x)\Phi_0(|x|)$, де $\Phi_0(x) = \int_0^x \varphi(t)dt$ – саме її значення найчастіше можна знайти в таблицях.

Теорема 5 (Гранична теорема Пуассона). Нехай кількість незалежних випробувань $n \to \infty$, причому $np_n \to \lambda$. Тоді $P_n(k) = C_n^k p_n{}^k q_n{}^{n-k} \to e^{-\lambda} \frac{\lambda^k}{k!}$ $(k = 0,1,2,\ldots,n,\ldots)$

Зауваження. Звісно, не можна вважати, що ймовірність успіху p змінюється в одній послідовності випробувань. Проте ми розглядаємо цю теорему у «схемі серій», тобто вважаємо що проводяться послідовні серії випробувань Бернуллі, і разом із зростанням кількості випробувань спадає ймовірність успіху, але так, що виконана умова $np_n \to \lambda$.

Доведення. Дійсно,

$$P_n(k) = \frac{n!}{k!(n-k)!} p_n{}^k q_n{}^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k! n^k (1-p_n)^k} (np_n)^k (1-p_n)^n =$$

$$= \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!(1-p_n)^k} (np_n)^k (1-p_n)^n$$

При $n \to \infty$, очевидно, $\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{(1-p_n)^k} \to 1$, $(np_n)^k \to \lambda^k$,

$(1-p_n)^n = \left((1-p_n)^{-\frac{1}{p_n}}\right)^{-np_n} \to e^{-\lambda}$, отже, при $n \to \infty$

$$P_n(k) \to e^{-\lambda} \frac{\lambda^k}{k!} = \pi_k(\lambda), \quad k = 0,1,2,\ldots \tag{3}$$

Граничні ймовірності $\pi_k(\lambda)$ мають властивість $\sum_{k=0}^{\infty} \pi_k(\lambda) = 1$ (яку легко перевірити) і називаються ***розподілом Пуассона з параметром*** λ. Значення цих ймовірностей також можна знайти в спеціальних таблицях. Теорема Пуассона справедлива при $p_n \cong \frac{\lambda}{n}, n \to \infty$, отже, при p_n близьких до нуля, тому теорему Пуассона часто називають законом ***рідкісних подій***. Досить точні наближення досягаються лише при $p_n \in (0,10)$. Для ймовірностей успіху поза межами рекомендованого інтервалу варто користатись наближенням, яке дає гранична теорема Муавра-Лапласа.

Розглянемо приклади застосування граничних теорем в схемі Бернуллі.

Приклад. Обчислити ймовірність того, що при 200 підкиданнях симетричної монети кількість випадінь герба відрізнятиметься від 100 не більше, ніж на 5.

Нехай ймовірність випадіння герба $p = 0,5$, кількість повторень експерименту $n = 200$, отже, значення $np = 100$ – досить велике, тому використовуватимемо граничну теорему Муавра-Лапласа.

Позначимо k – кількість випадінь герба при 200 підкиданнях монети. Знайдемо

$$P(|k-100| \le 5) = P(95 \le k \le 105) \approx \Phi\left(\frac{105-100}{\sqrt{200 \cdot 0,5 \cdot 0,5}}\right) - \Phi\left(\frac{95-100}{\sqrt{200 \cdot 0,5 \cdot 0,5}}\right) =$$

$$= \Phi\left(\frac{1}{\sqrt{2}}\right) - \Phi\left(-\frac{1}{\sqrt{2}}\right) = 2\Phi_0\left(\frac{\sqrt{2}}{2}\right) = 2\Phi_0(0,7778) \approx 0,56331$$

Приклад. Ймовірність виготовлення бракованої деталі на виробництві рівна $p = 0{,}005$. Яка ймовірність, що серед 10000 випадково обраних деталей кількість бракованих дорівнює 40? Не перевищує 40?

Оскільки в даному випадку $np = 50$, хоча й ймовірність p зовсім мала, знову використовуємо наближення Мавра-Лапласа. Тут

$$\sqrt{npq} = \sqrt{10000 \cdot 0{,}005 \cdot 0{,}995} \approx 7{,}05; \frac{k - np}{\sqrt{npq}} = \frac{40 - 50}{7{,}05} \approx -1{,}42,$$

тому $P_{10000}(40) \approx \frac{1}{7{,}05}\varphi(1{,}42) \approx \frac{0{,}1456}{7{,}05} \approx 0{,}0206$ Для відповіді на друге питання шукаємо ймовірність

$$P(0 \le k \le 40) = \sum_{k=0}^{40} P_{10000}(k) \approx \Phi\left(\frac{40-50}{7{,}05}\right) - \Phi\left(\frac{0-50}{7{,}05}\right) \approx$$

$$\approx 0{,}5 - \Phi_0(1{,}42) \approx 0{,}5 - 0{,}4222 = 0{,}0778$$

Приклад. Ймовірність влучити в ціль рівна $p = 0{,}001$. Для поразки цієї цілі потрібно не менше двох влучань. Знайти ймовірність враження цілі при 5000 пострілів.

Тут $np = 5$, тому в якості наближення використовуємо розподіл Пуассона. Знайдемо

$$P(k \ge 2) = \sum_{k=2}^{5000} P_{5000}(k) = 1 - \sum_{k=0}^{1} P_{5000}(k) = 1 - P_{5000}(0) - P_{5000}(1) \approx$$

$$\approx 1 - \pi_0(5) - \pi_1(5) = 1 - e^{-5} - 5e^{-5} \approx 0{,}9596$$

Приклад. Нехай в серії з n випробувань Бернуллі успіх стається з ймовірністю p. Оцінити ймовірність, що відхилення частоти успіху від його ймовірності не перевищує заданої величини ε.

Отже, маємо оцінити ймовірність:

$$P\left(\left|\frac{k}{n} - p\right| \le \varepsilon\right) = P\left(\left|\frac{k - np}{\sqrt{npq}}\right| \le \varepsilon\sqrt{\frac{n}{pq}}\right) \approx \Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right) = 2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Якщо $n \to \infty$, то $P\left(\left|\frac{k}{n} - p\right| \le \varepsilon\right) \to 1$. Цей результат відомий як ***теорема Бернуллі***, яка показує, що частота успіху в схемі Бернуллі із зростанням кількості випробувань відрізняється від ймовірності успіху менше від будь-якого малого числа ε з ймовірністю, як завгодно близькою до 1.

Лекція 4.Випадкові величини та їх розподіли.

1. Поняття випадкової величини.

Приклад 1. Розглянемо деякий стохастичний експеримент, наприклад, підкидання двох монет. Його простір елементарних подій $\Omega = \{\omega_1 = "ЦЦ", \omega_2 = "ГЦ", \omega_3 = "ЦГ", \omega_4 = "ГГ"\}$ складається із 4 рівноможливих елементарних подій. Спробуємо «виміряти» числом кожну з них. Одна з можливостей – поставити у відповідність кожній елементарній події $\omega_k\,(k = \overline{1,4})$ кількість випадінь герба. Таким чином, матимемо функцію $\xi : \Omega \to \Re$, значення

якої залежать від випадку. Така функція є ***випадковою величиною***. Її можна характеризувати ймовірностями, з якими вона набуває своїх значень: $P\{\xi(\omega)=0\}=P\{\xi(\omega)=2\}=1/4$, $P\{\xi(\omega)=1\}=1/2$.

Приклад 2. Розглянемо інший приклад – відбувається стрільба по мішені до першого влучного пострілу з ймовірністю p влучання в кожному окремому незалежному пострілі. Позначаючи влучання літерою «У», а промах – «Н», простір елементарних подій запишемо наступним чином:

$$\Omega=\left\{\omega_1="У",\omega_2="НУ",\omega_3="ННУ",\ldots,\omega_n="\underbrace{Н\ldots Н}_{n}У",\ldots\right\}.$$

Визначимо випадкову величину $\xi:\Omega\to\mathbf{R}$ як кількість пострілів до першого влучання. Тоді $P\{\xi(\omega)=k\}=P\{\omega_k\}=p(1-p)^{\mathrm{k}}\ \ \forall k=\overline{0,\infty}$. Ця випадкова величина визначена на нескінченому, проте знов таки дискретному просторі елементарних подій.

Приклад 3. Нехай точка випадковим чином кидається на відрізок $[a,b]$. В даному прикладі маємо неперервний простір елементарних подій $\Omega=\{\omega:\omega\in[a,b]\}$. Тоді випадкова величина, визначена як координата випадкової точки, $\xi:\Omega\to\mathbf{R}$ також має неперервний простір аргументів, при цьому, як відомо, (див. лекцію 1) $P\{\xi\in[\alpha,\beta]\}=\dfrac{\beta-\alpha}{b-a}$ $([\alpha,\beta]\subset[a,b])$.

Дамо нарешті точне означення випадкової величини. Нехай $(\Omega,\Im,P)$ – імовірнісний простір, де Ω – простір елементарних подій даного експерименту, $\Im$ – σ-алгебра випадкових (вимірних) подій, P – ймовірність випадкової події із $\Im$.

Означення 1. ***Випадковою величиною*** на імовірнісному просторі $(\Omega,\Im,P)$ називається функція $\xi:\Omega\to\mathbf{R}$, ***вимірна*** відносно σ-алгебри $\Im$, тобто така, що $\{\omega\in\Omega:\xi(\omega)<x\}\in\Im\ \ \forall x\in\mathbf{R}$.

Надалі будемо вважати фіксованим деякий імовірнісний простір $(\Omega,\Im,P)$.

Теорема 1. Нехай ξ – випадкова величина. Тоді $\{\omega\in\Omega:\xi(\omega)\le x\}$, $\{\omega\in\Omega:\xi(\omega)=x\}$, $\{\omega\in\Omega:\xi(\omega)>x\}$, $\{\omega\in\Omega:\xi(\omega)\ge x\}$, $\{\omega\in\Omega:a\le\xi(\omega)<b\}$ $\forall x,a,b\in\mathbf{R}$ є випадковими подіями.

Доведення. Для доведення необхідно показати, що всі зазначені вище множини елементарних подій належать σ-алгебра випадкових подій. Розглянемо, наприклад, множину

$\{\omega\in\Omega:\xi(\omega)\le x\}=\bigcap\limits_{n=1}^{\infty}\{\omega\in\Omega:\xi(\omega)<x+1/n\}\in\Im$, тобто це випадкова подія $\forall x\in\mathbf{R}$.

Тоді $\{\omega\in\Omega:\xi(\omega)=x\}=\{\omega\in\Omega:\xi(\omega)\le x\}\setminus\{\omega\in\Omega:\xi(\omega)<x\}\in\Im$. Якими б не були числа $a,b\in\mathbf{R}$: $\{\omega\in\Omega:a\le\xi(\omega)<b\}=\{\omega\in\Omega:\xi(\omega)<b\}\setminus\{\omega\in\Omega:\xi(\omega)<a\}\in\Im$. Інші твердження доводяться аналогічно.

Зауваження. Доведені властивості означають, що ймовірності цих подій цілком визначені.

Означення 2. ***Функцією розподілу*** випадкової величини $\xi:\Omega\to\mathbf{R}$ називається функція $F_\xi(x)=P\{\omega\in\Omega:\xi(\omega)<x\}\overset{def}{=}P\{\xi<x\}$ визначена $\forall x\in\mathbf{R}$.

Приклад 1*. Нехай ξ – випадкова величина, визначена у прикладі 1. Тоді

$$\{\omega\in\Omega:\xi(\omega)<x\}=\begin{cases}\varnothing, & x\le 0\\ \{\xi(\omega)=0\}=\omega_1, & 0<x\le 1\\ \{\xi(\omega)=0,\xi(\omega)=1\}=\omega_1\cup\omega_2\cup\omega_3, & 1<x\le 2\\ \{\xi(\omega)=0,\xi(\omega)=1,\xi(\omega)=2\}=\omega_1\cup\omega_2\cup\omega_3\cup\omega_4=\Omega, & x>2\end{cases}.$$

Отже, $$F_\xi(x)=P\{\xi<x\}=\begin{cases}0, & x\le 0\\ 1/4, & 0<x\le 1\\ 3/4, & 1<x\le 2\\ 1, & x>2\end{cases}.$$

Таким чином, для даної випадкової величини ξ її функція розподілу $F_\xi(x)$ – кусково-стала функція. Графік функції $F_\xi(x)$ на рисунку 1.

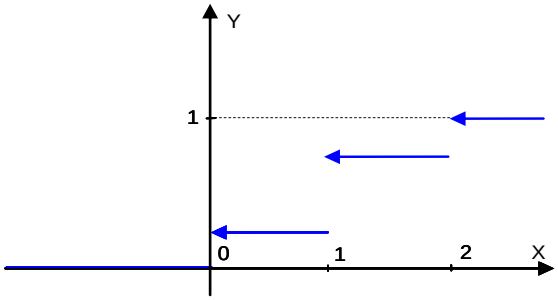


Рис.1

Приклад 3*. Нехай ξ – випадкова величина, визначена у прикладі 3. Тоді

$$\{\omega\in\Omega:\xi(\omega)<x\}=\begin{cases}\varnothing, & x\le a\\ \{\omega\in\Omega:\omega<x\}=[a,x), & a<x\le b\\ \Omega=[a,b], & x>b\end{cases}.$$

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

В даному прикладі функція $F_\xi(x)$ – неперервна. Її графік зображений на рисунку 2.

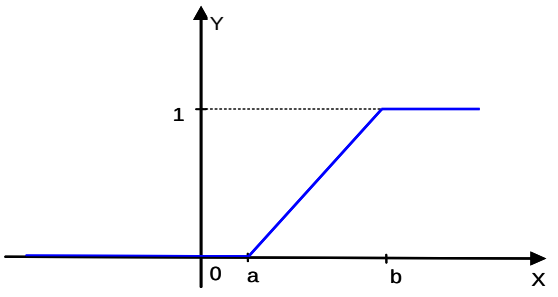


Рис. 2

Зауваження. Графіки функцій розподілу цих двох випадкових величин демонструють дещо схожу поведінку. (Яку саме?) Дослідимо детально властивості функції розподілу довільної випадкової величини ξ.

2. Властивості функції розподілу випадкової величини.

Теорема 2 (про властивості функції розподілу випадкової величини). Нехай ξ – випадкова величина на ймовірнісному просторі $(\Omega, \Im, P)$. $F_\xi(x)$ – її функція розподілу. Тоді:

1. $F_\xi(x)$ монотонно не спадає на всій числовій осі;
2. $\lim\limits_{x\to-\infty} F_\xi(x)=0,\ \lim\limits_{x\to+\infty} F_\xi(x)=1;$
3. $F_\xi(x)$ неперервна зліва, тобто $\lim\limits_{x\to x_0-0} F_\xi(x)=F_\xi(x_0);$
4. $P\{\omega\in\Omega: a\le\xi(\omega)<b\}=F(b)-F(a) \quad \forall a,b\in\mathbf{R}: a<b;$
5. $P\{\omega\in\Omega:\xi(\omega)\le x\}=F_\xi(x+0) \quad \forall x\in\mathbf{R}.$

Доведення.

1. Розглянемо довільні $x_1, x_2\in\mathbf{R}$, де $x_1<x_2$. Тоді $\{\omega\in\Omega:\xi(\omega)<x_1\}\subseteq\{\omega\in\Omega:\xi(\omega)<x_2\}$, тому $F_\xi(x_1)\le F_\xi(x_2)$, тобто $F_\xi(x)$ – монотонно не спадає на всій числовій осі.

2. Розглянемо довільну числову послідовність $\{x_n, n\ge1\}$, яка монотонно спадає до $-\infty$: $x_n\downarrow-\infty$ при $n\to\infty$. Введемо в розгляд події $A_n=\{\omega\in\Omega:\xi(\omega)<x_n\}$. Зрозуміло, що $A_{n+1}\subseteq A_n\,\forall n\ge1$. Згідно властивості неперервності ймовірності (теорема 5 з лекції 2) $P\left(\bigcap\limits_{n=1}^{\infty}A_n\right)=\lim\limits_{n\to\infty}P(A_n)=\lim\limits_{n\to\infty}F_\xi(x_n)=\lim\limits_{x\to-\infty}F_\xi(x)$ (Тут використано означення границі функції за Гейне). З іншого боку, $P\left(\bigcap\limits_{n=1}^{\infty}A_n\right)=P(\varnothing)=0$, отже, $\lim\limits_{x\to-\infty}F_\xi(x)=0$. Аналогічно можна обчислити і другу границю.

3. Розглянемо довільну числову послідовність $\{x_n, n\ge1\}$, яка монотонно зростає до точки x_0: $x_n\uparrow x_0$ при $n\to\infty$. Тоді $\{\omega\in\Omega:\xi(\omega)<x_0\}=\bigcup\limits_{n=1}^{\infty}\{\omega\in\Omega:\xi(\omega)<x_n\}$. Для монотонно зростаючої послідовності подій $A_n=\{\omega\in\Omega:\xi(\omega)<x_n\}$ з неперервності ймовірності випливає, що

$$F_\xi(x_0)=P\{\omega\in\Omega:\xi(\omega)<x_0\}=P\left(\bigcup_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}P(A_n)=\lim_{n\to\infty}F_\xi(x_n)=\lim_{x\to x_0-0}F_\xi(x).$$

4. Нехай $a<b$. Тоді $\{\omega\in\Omega: a\le\xi(\omega)<b\}=\{\omega\in\Omega:\xi(\omega)<b\}\setminus\{\omega\in\Omega:\xi(\omega)<a\}$, і оскільки перша подія є наслідком другої, то з теореми 3 з лекції 2 випливає, що $P\{\omega\in\Omega: a\le\xi(\omega)<b\}=P\{\omega\in\Omega:\xi(\omega)<b\}-P\{\omega\in\Omega:\xi(\omega)<a\}=F(b)-F(a).$

5. Нехай довільна числова послідовність $\{x_n, n\ge1\}$ монотонно спадає до деякого числа x: $x_n\downarrow x$ при $n\to\infty$. Тоді $\{\omega\in\Omega:\xi(\omega)\le x\}=\bigcap\limits_{n=1}^{\infty}\{\omega\in\Omega:\xi(\omega)<x_n\}$. Для монотонно спадаючої послідовності подій $\{\omega\in\Omega:\xi(\omega)<x_n\}$ справедливо, що

$$P\{\omega\in\Omega:\xi(\omega)\le x\}=P\left(\bigcap_{n=1}^{\infty}\{\omega\in\Omega:\xi(\omega)<x_n\}\right)=\lim_{n\to\infty}P(\{\omega\in\Omega:\xi(\omega)<x_n\})=F_\xi(x+0)$$

Наслідок. $\forall x \in \mathbf{R}:\ P\{\omega \in \Omega : \xi(\omega) = x\} = F_\xi(x+0) - F_\xi(x)$, отже, якщо функція розподілу деякої випадкової величини неперервна, то $\forall x \in \mathbf{R}:\ P\{\omega \in \Omega : \xi(\omega) = x\} = 0$.

Зауваження. Виявляється, що визначальними для функції розподілу є властивості 1 – 3. Іншими словами, якщо деяка функція $F(x), x \in \mathbf{R}$ має властивості 1 – 3, то можна визначити імовірнісний простір $(\Omega, \Im, P)$ та випадкову величину ξ на ньому такі, що $F_\xi(x) \equiv F(x)$.

3. Дискретні випадкові величини.

Означення 3. Випадкова величина $\xi : \Omega \to \mathbf{R}$ називається ***дискретною***, якщо множина її значень скінчена або злічена.

Теорема 3. Дискретна випадкова величина повністю характеризується набором своїх значень та ймовірностями, з якими ці значення набуваються.

Доведення. Дійсно, нехай випадкова величина ξ набуває значень $x_1, x_2, \ldots$, причому $P\{\omega \in \Omega : \xi(\omega) = x_k\} = p_k \geq 0$, де $\sum_k p_k = 1$.

Тоді $F_\xi(x) = P\{\omega \in \Omega : \xi(\omega) < x\} = \sum_{k: x_k < x} P\{\omega \in \Omega : \xi(\omega) = x_k\} = \sum_{k: x_k < x} p_k$.

Означення 4. Набір ймовірностей $p_k, k \geq 1$ називається ***розподілом*** дискретної випадкової величини ξ.

Приклад 4. Нехай n – кількість незалежних випробувань в схемі Бернуллі з ймовірністю успіху p в кожному випробуванні. Нехай величина ξ визначає кількість успішних випробувань. Тоді, як відомо, $P\{\xi = k\} = C_n^k p^k q^{n-k}$, $k = \overline{0, n}$. Нагадаємо тут, що ця випадкова величина має ***біноміальний розподіл*** з параметрами n та p.

Приклад 2*. Повернемось до випадкової величини, розглянутої в прикладі 2. Ця дискретна випадкова величина має ***геометричний розподіл*** з параметром $p > 0$ (така випадкова величина вже згадувалась в лекції 3). Геометрично розподілена випадкова величина має властивість ***відсутності післядії***. А саме: $P\{\xi = n + m / \xi \geq n\} = P\{\xi = m\}\ \forall m \geq 0, n \geq 1$ (доведіть це самостійно). Можна дати наступну ілюстрацію цієї властивості. Нехай ξ означає випадкову тривалість телефонної розмови, причому на початку кожної хвилини розмови вона або продовжується ще на хвилину з ймовірністю p, або припиняється із ймовірністю $1 - p$. Тоді ймовірність того, що розмова триватиме ще m хвилин, якщо вона вже триває не менше ніж n хвилин збігається із ймовірністю, що розмова від початку триватиме рівно m хвилин незалежно від n. Крім того, можна показати, що в класі дискретних розподілів лише геометричний має таку властивість.

Приклад 5. Випадкова величина ξ має розподіл Пуассона з параметром $\lambda > 0$ (ця випадкова величина також згадувалась в лекції 3), якщо вона має своїми значеннями всі цілі невід'ємні числа, причому $P\{\xi = n\} = \dfrac{\lambda^n e^{-\lambda}}{n!}\ \forall n \geq 0$.

Завдання для самостійної роботи. Нехай випадкова величина ξ – це кількість успіхів в серії із 3 незалежних випробувань Бернуллі з ймовірністю успіху $p = 2/3$. Побудувати графік функції розподілу даної випадкової величини.

4. Абсолютно неперервні випадкові величини.

Означення 5. Випадкова величина $\xi:\Omega\to\mathbf{R}$ називається ***абсолютно неперервною***, якщо $\forall x\in\mathbf{R}$ значення її функції розподілу $F_\xi(x)$ може бути представлене у вигляді інтегралу: $F_\xi(x)=\int\limits_{-\infty}^{x}p_\xi(u)du$, де функція $p_\xi(x), x\in(-\infty,\infty)$ називається ***щільністю розподілу*** випадкової величини ξ.

Теорема 4 (про властивості щільності розподілу випадкової величини).

1. $p_\xi(x)$ – невід'ємна: $p_\xi(x)\geq 0\ \forall x\in\mathbf{R}$;
2. $\int\limits_{-\infty}^{+\infty}p_\xi(u)du=1$;
3. $P\{\omega\in\Omega: a\leq\xi(\omega)<b\}=\int\limits_{a}^{b}p_\xi(u)du\ \ \forall a,b\in\mathbf{R}: a<b$;

Доведення.

1. З існування щільності випливає, що $F_\xi(x)$ – диференційована функція і $F_\xi{}'(x)=p_\xi(x)$. А оскільки $F_\xi(x)$ монотонно не спадає, то $p_\xi(x)\geq 0\ \ \forall x\in\mathbf{R}$.

2. Ця властивість випливає з того, що $\lim\limits_{x\to+\infty}F_\xi(x)=\int\limits_{-\infty}^{+\infty}p_\xi(u)du=1.$

3. Це прямий наслідок теореми Ньютона-Лейбніца. Дана властивість означає, що ймовірність $P\{a\leq\xi<b\}$ рівна площі фігури під графіком щільності випадкової величини на відрізку $[a,b]$.

Наслідок. З даної теореми та теореми про середнє значення інтегралу випливає, що $P\{x\leq\xi(\omega)<x+\Delta x\}=p_\xi(x)\cdot\Delta x+o(\Delta x),\ \Delta x\to 0$ для будь-якої точки неперервності функції $p_\xi(x)$. Саме для абсолютно неперервної випадкової величини ймовірність набути будь-яке значення дорівнює нулю, проте, взагалі кажучи, не є неможливою подією.

Приклад 3**. Повернемось до випадкової величини з прикладу 3. Якщо знайти похідну від її функції розподілу, одержимо: $F_\xi{}'(x)=p_\xi(x)=\begin{cases}\dfrac{1}{b-a}, & x\in(a,b)\\ 0, & x\notin(a,b)\end{cases}$

Означення 6. Випадкова величина, що має таку щільність розподілу, називається ***рівномірно розподіленою*** на відрізку $[a,b]$.

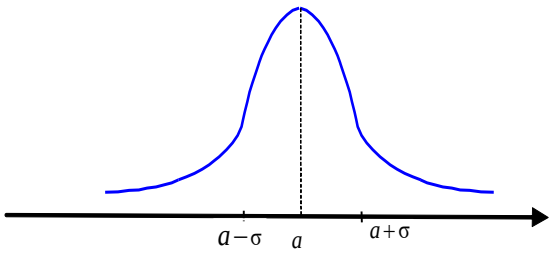


Рис.3

Означення 7. Випадкова величина, яка має щільність розподілу, визначену формулою:

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

називається ***нормально розподіленою*** з параметрами a та σ^2. Для

нормального розподілу (його ще називають гауссовим) використовують позначення $N(a,\sigma^2)$. Графік функції нормального розподілу схематично представлений на рис.3.

Зауваження. Пригадайте функцію $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ із граничної теореми Муавра-Лапласа. Це була щільність розподілу випадкової величини із нормальним $N(0,1)$ розподілом (нормальний розподіл з такими параметрами називається стандартним нормальним розподілом).

Означення 8. Випадкова величина, яка має щільність розподілу, визначену формулою $p(x)=\begin{cases}\lambda e^{-\lambda x}, x>0\\ 0, \quad x\le 0\end{cases}$, називається ***показниково розподіленою*** з параметром $\lambda>0$. Графік цієї щільності наведений на рис. 4.

Її функція розподілу визначається формулою: $F(x)=\begin{cases}1-e^{-\lambda x}, x>0\\ 0, \quad x\le 0\end{cases}$

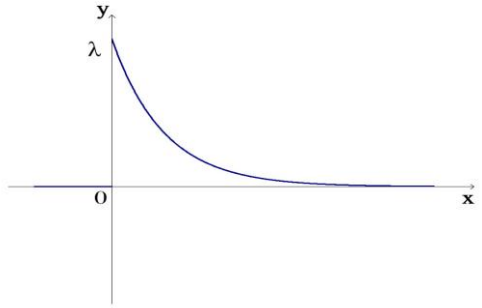


Рис. 4

Надалі часто виникатиме задача відшукання розподілу деякої функції від випадкової величини. Тому перш за все треба з'ясувати, які саме функції від випадкових величин в свою чергу будуть випадковими величинами. Відповідь на це питання дає наступна теорема, яку ми змушені будемо лишити без доведення.

Теорема. Нехай $(\Omega,\Im,P)$ – імовірнісний простір і ξ – випадкова величина на ньому. Нехай $f:\mathbf{R}\to\mathbf{R}$ – борелівська функція (тобто, вимірна відносно σ-алгебри борелівських множин $\mathsf{B}(\mathbf{R})$ на прямій $\mathbf{R}$: $\{x\in\mathbf{R}: f(x)<c\}\in\mathsf{B}(\mathbf{R})\ \forall c\in\mathbf{R}$). Тоді $f(\xi)$ – випадкова величина на $(\Omega,\Im,P)$.

Зауваження. Перевіряти, чи є функція борелівською – не проста справа, проте принаймні всі неперервні функції є борелівськими, і це дає нам досить широкі можливості по вивченню розподілів нових випадкових величин, які є функціями від уже відомих.

Приклад 6. Нехай ξ – рівномірно розподілена на відрізку [0,1] випадкова величина. Визначити розподіл випадкової величини $\eta=\ln 1/\xi$.

Зрозуміло, що $\eta=\ln 1/\xi$ так само, як і ξ, буде неперервною випадковою величиною в області визначення. І хоча характеризувати її треба щільністю розподілу, шукатимемо спочатку її функцію розподілу:

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

Отже, $F_\eta(x)=P(\xi>e^{-x})=1-P(\xi\le e^{-x})=1-F_\xi(e^{-x})=\begin{cases}1-e^{-x}, \ при \ \ x>0\\ 0, \ при \ \ x\le 0\end{cases}$, тому що

для рівномірно розподіленої величини ξ $P(\xi \le x) = F_\xi(x) = x$ при $0 < x < 1$. Одержаний результат означає, що випадкова величина η має показниковий розподіл з параметром $\lambda = 1$. Її щільність рівна $p_\eta(x) = F'_\eta(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & x \le 0 \end{cases}$.

Приклад 7. Нехай ξ – випадкова величина із неперервною функцією розподілу $F(x)$. Визначити розподіл випадкової величини $\eta = F(\xi)$.

Знайдемо функцію розподілу випадкової величини η: $F_\eta(x) = P(\eta < x) = P(F(\xi) < x)$. Оскільки $F(x)$ – це функція розподілу, то $F_\eta(x)$ рівна 0 при $x \le 0$ і $F_\eta(x)$ рівна 1 при $x > 1$. Нехай $0 < x \le 1$. Розглянемо точку $z \in \mathbf{R}$, для якої $F(z) = x$, тобто $z = F^{-1}(x)$ (Із-за неперервності та монотонності функції $F(x)$ така точка визначається однозначно). Тоді $F_\eta(x) = P(F(\xi) < x) = P\left(\xi < F^{-1}(x)\right) = P(\xi < z) = F(z) = F\left(F^{-1}(x)\right) = x$. Отже, при $0 < x \le 1$ щільність розподілу випадкової величини η рівна 1, тобто η – рівномірно розподілена випадкова величина на $[0,1]$.

Зауваження 1. Дискретними та неперервними випадковими величинами далеко не вичерпуються всі можливі розподіли. Крім випадкових величин, які в деякому інтервалі ведуть себе як дискретні, а в іншому – як неперервні, існують ще так звані сингулярні розподіли. Їх функція розподілу може бути неперервна, проте похідна від неї рівна нулю в усіх точках прямої $\mathbf{R}$, крім множини точок міри 0. Надалі такі розподіли не розглядатимуться, оскільки вони вимагають для свого вивчення більш потужного математичного апарату, ніж той, яким навіть в ідеалі має володіти студент-радіофізик.

Зауваження 2. Нехай $y = f(x)$ – взаємно-однозначне відображення із $\mathbf{R}$ в $\mathbf{R}$. Тоді існує обернене відображення $f^{-1} \overset{def}{=} g$, тобто $x = g(y)$. Якщо випадкова величина ξ має щільність розподілу $p_\xi(x)$, то випадкова величина $\eta = f(\xi)$ також має щільність розподілу, яка визначається формулою: $p_\eta(y) = p_\xi(g(y)) \cdot |g'(y)|$.

Лекція 5. Випадкові вектори (Багатовимірні розподіли).

1. Поняття про випадковий вектор та його розподіл.

Нехай $(\Omega, \Im, P)$ – імовірнісний простір, де Ω – простір елементарних подій даного експерименту, $\Im$ – σ-алгебра випадкових (вимірних) подій, P – ймовірність випадкової події із $\Im$. На даному імовірнісному просторі розглянемо сукупність n випадкових величин $\xi_1 = \xi_1(\omega), \xi_2 = \xi_2(\omega), \ldots, \xi_n = \xi_n(\omega)$.

Означення 1. ***Випадковим вектором*** або n-вимірною випадковою величиною на імовірнісному просторі $(\Omega, \Im, P)$ називається вектор $\overline{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$.

Очевидно, що множина

$$\{\omega \in \Omega : \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\} \overset{def}{=} \{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n\}$$

є випадковою подією при всіх значеннях $x_1, x_2, \ldots, x_n \in \mathbf{R}$, отже, визначена функція n змінних $F_{\xi_1, \xi_2, \ldots, \xi_n}(x_1, x_2, \ldots, x_n) = P\{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n\}$, яка називається

n*-вимірною функцією розподілу** випадкового вектора $\bar{\xi}$ або ***сумісним розподілом випадкових величин $\xi_1, \xi_2, \ldots, \xi_n$. Наступні властивості n-вимірних функцій розподілу доведіть самостійно.

Теорема 1. (Властивості багатовимірних розподілів).

1. $F_{\xi_1,\xi_2,\ldots,\xi_n}(x_1, x_2, \ldots, x_n)$ – монотонно неспадна функція по кожному своєму аргументу, тобто: $F_{\xi_1,\ldots,\xi_k,\ldots,\xi_n}(x_1, \ldots, x_k', \ldots, x_n) \le F_{\xi_1,\ldots,\xi_k,\ldots,\xi_n}(x_1, \ldots, x_k'', \ldots, x_n)$ при довільних $x_k' < x_k'' \; \forall x_i', i = 1, \ldots, k-1, k+1, \ldots, n$.
2. $F_{\xi_1,\xi_2,\ldots,\xi_n}(x_1, x_2, \ldots, x_n)$ – неперервна зліва функція по кожному своєму аргументу, тобто:
   $\lim\limits_{x_k \to x_0 \; 0} F_{\xi_1,\ldots,\xi_k,\ldots,\xi_n}(x_1, \ldots, x_k, \ldots, x_n) = F_{\xi_1,\ldots,\xi_k,\ldots,\xi_n}(x_1, \ldots, x_0, \ldots, x_n) \;\; \forall x_k, k = \overline{1,n}$.
3. $\lim\limits_{x_k \to -\infty} F_{\xi_1,\ldots,\xi_k,\ldots,\xi_n}(x_1, \ldots, x_k, \ldots, x_n) = 0 \;\; \forall x_k, k = \overline{1,n}$.
4. $\forall x_k, k = \overline{1,n}$:
   $\lim\limits_{x_k \to +\infty} F_{\xi_1,\ldots,\xi_k,\ldots,\xi_n}(x_1, \ldots, x_k, \ldots, x_n) = F_{\xi_1,\ldots,\xi_{k-1},\xi_{k+1},\ldots,\xi_n}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$
   тобто гранична функція є $(n-1)$-вимірною функцією розподілу випадкового вектора $(\xi_1, \ldots, \xi_{k-1}, \xi_{k+1}, \ldots, \xi_n)$.
5. $\lim\limits_{\substack{x_1 \to +\infty, \\ \vdots \\ x_n \to +\infty}} F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = 1$.

Зокрема, при $n = 2$ сумісний розподіл випадкового вектора (ξ, η) $F_{\xi,\eta}(x, y) = P\{\xi < x, \; \eta < y\}$ дозволяє визначити одновимірні (маргінальні) розподіли як границі: $F_\xi(x) = \lim\limits_{y \to +\infty} F_{\xi,\eta}(x, y)$, $F_\eta(y) = \lim\limits_{x \to +\infty} F_{\xi,\eta}(x, y)$, проте невірно, що одновимірні розподіли компонент вектора дозволяють визначити сумісний розподіл, тобто відомостей про поведінку компонент випадкового вектора, взагалі кажучи, недостатньо для визначення поведінки всього вектора.

Якщо ξ, η – дискретні випадкові величини, причому ξ набуває значень $x_1, x_2, \ldots, x_m, \ldots$, а η набуває значень $y_1, y_2, \ldots, y_n, \ldots$,то розподіл випадкового вектора (ξ, η) часто задають у вигляді таблиці:

$\xi \setminus \eta$	y_1	y_2	…	y_n	…
x_1	p_{11}	p_{12}	…	p_{1n}	…
x_2	p_{21}	p_{22}	…	p_{2n}	…
⋮	⋮	⋮		⋮	
x_m	p_{m1}	p_{m2}	…	p_{mn}	…
⋮					

Тут $p_{mn} = P(\xi = x_m, \eta = y_n)$, $m, n = \overline{1, \infty}$, отже, можна визначити розподіли випадкових величин ξ, η:

$$P(\xi = x_m) \overset{def}{=} p_m^\xi = \sum_{n=1}^{\infty} p_{mn}, \; \forall m = \overline{1, \infty}, \; P(\eta = y_n) \overset{def}{=} p_n^\eta = \sum_{m=1}^{\infty} p_{mn}, \; \forall n = \overline{1, \infty},$$

тобто суми рядків таблиці сумісного розподілу задають розподіл випадкової величини ξ, а суми стовпчиків таблиці – розподіл випадкової величини η.

Означення 2. Випадковий вектор (ξ,η) називається ***абсолютно неперервним (неперервним)***, якщо при всіх $x,y\in\mathbf{R}$ сумісна функція розподілу допускає інтегральне представлення:

$$F_{\xi,\eta}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p_{\xi,\eta}(u,v)dudv. \qquad (1)$$

Функція $p_{\xi,\eta}(x,y)$ називається ***сумісною щільністю*** розподілу випадкових величин ξ,η. Очевидно, що у всіх точках неперервності функції $p_{\xi,\eta}(x,y)$ має місце рівність $p_{\xi,\eta}(x,y)=\dfrac{\partial^2F(x,y)}{\partial x\partial y}$

Очевидні властивості сумісної щільності містяться у наступній теоремі.

Теорема 2. (Властивості сумісної щільності розподілу двох випадкових величин).

1. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p_{\xi,\eta}(u,v)dudv=1$ (2)
2. $p_{\xi,\eta}(x,y)\geq 0\ \forall x,y\in\mathbf{R}$
3. $P((\xi,\eta)\in G)=\iint_{G}p_{\xi,\eta}(u,v)dudv$ для довільної області $G\subseteq\mathbf{R}^2$.

Знаючи сумісну щільність розподілу випадкового вектору (ξ,η), можна визначити щільність кожної з випадкових величин, їх ще називають маргінальними розподілами, на відміну від сумісного (обернене невірно - маргінальні щільності, взагалі кажучи, не визначають сумісну щільність):

$F_\xi(x)=\lim_{y\to+\infty}F_{\xi,\eta}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{+\infty}p_{\xi,\eta}(u,v)dudv$, звідки $p_\xi(x)=F_\xi'(x)=\int_{-\infty}^{+\infty}p_{\xi,\eta}(x,v)dv$.

Аналогічно, $p_\eta(y)=\int_{-\infty}^{+\infty}p_{\xi,\eta}(u,y)du$

Приклад 1. Випадковий вектор (ξ,η) називається ***рівномірно розподіленим в області*** $G\subseteq\mathbf{R}^2$, якщо сумісна щільність розподілу його компонент визначається рівністю: $p_{\xi,\eta}(x,y)=\begin{cases}c, & (x,y)\in G\\ 0, & (x,y)\ \ G\end{cases}$ (3)

Для визначення константи c скористаємось властивістю (2) сумісної щільності розподілу. Нехай $S(G)$ – площа області G, тоді маємо: $1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p_{\xi,\eta}(u,v)dudv=\iint_{G}cdudv=cS(G)$, звідки $c=\dfrac{1}{S(G)}$. Отже, $P((\xi,\eta)\in D)=\iint_{D}p_{\zeta,\eta}(u,v)dudv=\dfrac{1}{S(G)}\iint_{D}dudv=\dfrac{S(D)}{S(G)}$, тобто поняття геометричної ймовірності пов'язане саме з рівномірним розподілом в області.

Приклад 2. Випадковий вектор (ξ,η) називається ***нормально розподіленим*** з параметрами $a_1,a_2,{\sigma_1}^2,{\sigma_2}^2$ та r, якщо сумісна щільність розподілу його компонент визначається формулою:

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{(x-a_1)^2}{{\sigma_1}^2}+\frac{(y-a_2)^2}{{\sigma_2}^2}-\frac{2r(x-a_1)(y-a_2)}{\sigma_1\sigma_2}\right)\right\} \quad (4)$$

Завдання для самостійної роботи. Покажіть, що при $r=0$ випадкові величини ξ та η мають нормальні розподіли $N(a_1,{\sigma_1}^2)$ та $N(a_2,{\sigma_2}^2)$ відповідно.

Означення 3. Випадкові величини ξ та η будемо називати ***незалежними***, якщо

$$P\{\xi<x,\eta<y\}=P\{\xi<x\}\cdot P\{\eta<y\}\ \forall x,y\in\mathbf{R} \quad (5)$$

Іншими словами, випадкові величини незалежні тоді і тільки тоді, коли сумісна функція їх розподілу є добутком їх функцій розподілу:

$$F_{\xi,\eta}(x,y)=F_\xi(x)\cdot F_\eta(y)\ \forall x,y\in\mathbf{R}$$

Зауваження 1. Справедливе і більш загальне твердження: випадкові величини ξ та η незалежні тоді і тільки тоді, коли події $P\{\xi\in B_1\}$ та $P\{\eta\in B_2\}$ – незалежні для будь-яких борелівських множин B_1 та B_2.

Зауваження 2. Поняття незалежності, звісно, природно розповсюджується на довільну кількість випадкових величин.

Зауваження 3. Якщо ξ та η – дискретні випадкові величини з розподілами $P(\xi=x_m)=p_m^\xi,\ m=\overline{1,\infty}$ та $P(\eta=y_n)=p_n^\eta,\ n=\overline{1,\infty}$, то незалежні вони тоді і тільки тоді, коли $p_{mn}=P(\xi=x_m,\eta=y_n)=p_m^\xi\cdot p_n^\eta\ \forall m,n\geq 1$

Зауваження 4. Якщо ξ та η – неперервні випадкові величини, що мають щільності розподілів $p_\xi(x)$ та $p_\eta(y)$ відповідно, то незалежні вони тоді і тільки тоді, коли їх сумісна щільність розподілу є добутком їх щільностей: $p_{\xi,\eta}(x,y)=p_\xi(x)\cdot p_\eta(y)\ \forall x,y\in\mathbf{R}$

Приклад 3. (Розподіл суми двох випадкових величин). Нехай ξ та η – випадкові величини із сумісною щільністю розподілу $p_{\xi,\eta}(x,y)$. Знайдемо розподіл випадкової величини $\varsigma=\xi+\eta$.

Отже, $F_\varsigma(x)=P(\varsigma<x)=P(\xi+\eta<x)=P((\xi,\eta)\in D_x)$,

де $D_x=\{(u,v)\in\mathbf{R}^2:u+v<x\}$. Тоді $F_\varsigma(x)=\iint\limits_{D_x}p_{\xi,\eta}(u,v)dudv=\int\limits_{-\infty}^{+\infty}du\int\limits_{-\infty}^{x-u}p_{\xi,\eta}(u,v)dv$,

звідки $$p_\varsigma(x)=F'_\varsigma(x)=\int\limits_{-\infty}^{+\infty}p_{\xi,\eta}(u,x-u)du \quad (6)$$

Якщо випадкові величини ξ та η ***незалежні***, то щільність розподілу їх суми $\varsigma=\xi+\eta$ є ***згорткою*** щільностей доданків: $p_\varsigma(x)=\int\limits_{-\infty}^{+\infty}p_\xi(u)p_\eta(x-u)du$. (7)

Використаємо формулу (7), щоб знайти розподіл ***суми двох незалежних випадкових величин рівномірно розподілених відповідно на відрізках*** [-1,0] ***та*** [0,1]. Отже, маємо: $p_\xi(x)=\begin{cases}1, & x\in[-1,0]\\ 0, & x\notin[-1,0]\end{cases}$ та $p_\eta(y)=\begin{cases}1, & y\in[0,1]\\ 0, & y\notin[0,1]\end{cases}$, $p_{\xi,\eta}(x,y)=p_\xi(x)\cdot p_\eta(y)\ \forall x,y\in\mathbf{R}$. Тоді для випадкової величини $\varsigma=\xi+\eta$ одержимо:

$$p_\zeta(x)=\int_{-\infty}^{+\infty}p_\xi(u)p_\eta(x-u)du=\int_{-1}^{0}p_\eta(x-u)du=\begin{cases}0, & x\in(-\infty,-1)\\ \int_{-1}^{x}du, & x\in[-1,0]\\ \int_{x-1}^{0}du, & x\in(0,1]\\ 0, & x\in(1,+\infty)\end{cases}=\begin{cases}x+1, & x\in[-1,0]\\ -x+1, & x\in[0,1]\\ 0, & x\notin[-1,1]\end{cases}$$

Назва цього розподілу – трикутний на відрізку [-1,1] стає зрозумілою, якщо подивитись на графік щільності цього розподілу на рис. 1:

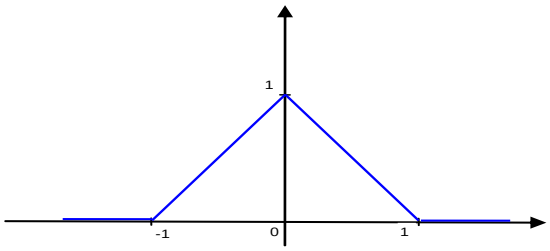


Рис. 1

Приклад 4. (χ^2-квадрат розподіл з n степенями свободи)

Нехай $\xi_1,\xi_2,\ldots,\xi_n$ – незалежні випадкові величини зі стандартними нормальними розподілами $N(0,1)$. Розглянемо випадкову величину $\eta=\xi_1{}^2+\xi_2{}^2+\cdots+\xi_n{}^2$, яка задає квадрат довжини випадкового вектора з компонентами $\xi_1,\xi_2,\ldots,\xi_n$. Можна показати, що щільність розподілу випадкової величини η рівна

$$p_\eta(x)=\begin{cases}0, & x\le 0\\ \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, & x>0\end{cases}\qquad(7)$$

χ^2-квадрат розподіл має широке застосування у математичній статистиці, зокрема для перевірки гіпотези про узгодженість емпіричних даних з деякою гіпотетичною функцією розподілу.

Приклад 6. (Γ-розподіл з параметрами α та β)

Випадкова величина має Γ розподіл з параметрами α та β, якщо її щільність задається формулою:

$$p_\eta(x)=\begin{cases}0, & x\le 0\\ \dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}, & x>0\end{cases}\qquad(8)$$

Цей розподіл є досить універсальним – при $\alpha=0$ він збігається із показниковим розподілом, при $\alpha=\frac{n}{2},\beta=\frac{1}{2}$ – з χ^2-квадрат розподілом з n степенями свободи.

Теорема 3. (Формула зміни сумісної щільності розподілу при заміні змінних).

Нехай $\mathbf{y}=\mathbf{f}(\mathbf{x})$ – взаємно-однозначне відображення $\mathbf{R}^n$ в $\mathbf{R}^n$, а $\mathbf{x}=\mathbf{g}(\mathbf{y})$ є оберненим відображенням. Нехай $\boldsymbol{\xi}=(\xi_1,\xi_2,\ldots,\xi_n)$ – випадковий вектор, а $\eta_1=f_1(\xi_1,\xi_2,\ldots,\xi_n),\ldots,\eta_n=f_n(\xi_1,\xi_2,\ldots,\xi_n)$. Довести, що щільність

$p_{(\eta_1,\eta_2,\ldots,\eta_n)}(y_1, y_2, \ldots, y_n)$ розподілу випадкового вектора $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_n)$ дорівнює:

$$p_{(\eta_1,\eta_2,\ldots,\eta_n)}(y_1, y_2, \ldots, y_n) =$$

$$= p_{(\xi_1,\xi_2,\ldots,\xi_n)}(g_1(y_1, y_2, \ldots, y_n), g_2(y_1, y_2, \ldots, y_n), \ldots, g_n(y_1, y_2, \ldots, y_n)) \left| \frac{D(x_1, x_2, \ldots, x_n)}{D(y_1, y_2, \ldots, y_n)} \right|,$$

де $\left| \frac{D(x_1, x_2, \ldots, x_n)}{D(y_1, y_2, \ldots, y_n)} \right|$ – якобіан перетворення $\mathbf{x} = \mathbf{g}(\mathbf{y})$.

Зауваження. Відповідна формула для одновимірного випадку була наведена в попередній лекції: $p_\eta(y) = p_\xi(g(y)) \cdot |g'(y)|$

Лекція 6. Числові характеристики випадкових величин.

Нехай ξ – випадкова величина на даному імовірнісному просторі $(\Omega, \Im, P)$.

1. Математичне сподівання випадкової величини та його властивості.

Означення 1. ***Математичним сподіванням*** або ***середнім*** випадкової величини ξ називається число $E\xi$, яке визначається формулою

$$E\xi = \sum_{k=1}^{\infty} x_k p_k \quad , \qquad (1)$$

якщо ξ – дискретна випадкова величина, що набуває значень $x_1, \ldots, x_k, \ldots$ із ймовірностями $p_1, \ldots, p_k, \ldots$ тобто $P\{\xi = x_k\} = p_k$, $k = 1,2,\ldots$;

або формулою

$$E\xi = \int_{-\infty}^{+\infty} xp(x)dx, \qquad (2)$$

якщо ξ – неперервна випадкова величина із щільністю розподілу $p(x)$.

Зауваження 1. Формули (1) та (2) мають сенс лише тоді, коли ряд (1) або невласний інтеграл (2) абсолютно збігаються. Тоді випадкова величина ξ називається ***інтегрованою***.

Зауваження 2. Число $E\xi$ ще називають ***першим моментом*** випадкової величини ξ.

Зауваження 3. Використовується також позначення $M\xi$ для математичного сподівання. У фізиці, оскільки літера «Е» означає енергію, для середнього випадкової величини використовують кутові дужки: $\langle \xi \rangle$.

Зауваження 4. Точною формулою для визначення $M\xi$ є інтеграла Лебега-Стілтьєса:

$$E\xi = \int_{-\infty}^{+\infty} x dF_\xi(x) \, , \qquad (3)$$

де $F_\xi(x)$ – функція розподілу випадкової величини ξ.

Наступні приклади пояснюють зміст терміну «середнє».

Приклад 1. Нехай ξ – рівномірно розподілена на відрізку $[a,b]$ випадкова величина. Тоді $E\xi = \int_{-\infty}^{+\infty} xp(x)dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$. В даному разі середнім виявилась середина відрізку $[a,b]$.

Приклад 2. Нехай ξ – дискретна випадкова величина, яка набуває n рівноімовірних значень $x_1, x_2, \ldots, x_n$. Згідно формулі (1), маємо $E\xi = \sum_{k=1}^{n} x_k p_k = \sum_{k=1}^{n} x_k \cdot \frac{1}{n} = \sum_{k=1}^{n} x_k \Big/ n$. В даному разі математичне сподівання – це середнє арифметичне значень дискретної випадкової величини ξ.

Означення 2. Випадкова величина, математичне сподівання якої рівне нулю, називається ***центрованою***.

Зауваження. Математичне сподівання має і фізичний зміст: якщо набір мас $p_1, \ldots, p_k, \ldots$ деякої системи мас знаходиться в точках прямої з координатами $x_1, \ldots, x_k, \ldots,$ або розподілений по прямій із щільністю розподілу $p(x)$, то центр ваги цієї системи знаходиться в точці $E\xi$.

Теорема 1. (Властивості математичного сподівання).

1. Якщо $\xi = c(const)$, то $E\xi = c$.
2. Якщо ξ – інтегрована невід'ємна випадкова величина, то $E\xi \geq 0$.
3. Якщо ξ – інтегрована випадкова величина, то випадкова величина $c\xi$, де c є довільною константою, інтегрована також, причому $E(c\xi) = cE\xi$.
4. Якщо ξ та η – інтегровані випадкові величини, то $\xi + \eta$ – також інтегрована, причому $E(\xi + \eta) = E\xi + E\eta$.
5. Для кожної борелівської функції $g(x), x \in \mathbf{R}$ та будь-якої інтегрованої випадкової величини ξ існує $Eg(\xi)$, причому, якщо ξ – дискретна випадкова величина, що набуває значень $x_1, \ldots, x_k, \ldots$ із ймовірностями $p_1, \ldots, p_k, \ldots : P\{\xi = x_k\} = p_k, k = 1,2,\ldots,$ то $Eg(\xi) = \sum_{k=1}^{\infty} g(x_k) p_k$, (4)

   а якщо ξ – неперервна випадкова величина із щільністю розподілу $p(x)$, то

   $$Eg(\xi) = \int_{-\infty}^{+\infty} g(x) p(x) dx . \quad (4')$$

6. Якщо ξ та η – ***незалежні*** інтегровані випадкові величини, то існує

   $$E(\xi - \eta) = E\xi - E\eta, \quad (5)$$

Доведення. Властивості **1. – 3.** неважко довести самостійно. Доведемо деякі наступні властивості (деякі лише для випадків дискретних випадкових величин).

4. Нехай ξ та η – дискретні випадкові величини для яких існують $E\xi$ та $E\eta$. Припустимо, що ξ набуває значень $x_1, \ldots, x_k, \ldots$ із ймовірностями $p_1^{\xi}, \ldots, p_k^{\xi}, \ldots : P\{\xi = x_k\} = p_k^{\xi}, k = 1,2,\ldots,$ а η набуває значень $y_1, \ldots, y_k, \ldots$ із ймовірностями $p_1^{\eta}, \ldots, p_k^{\eta}, \ldots : P\{\eta = y_k\} = p_k^{\eta}, k = 1,2,\ldots.$ Тоді випадкова величина $\xi + \eta$ набуває значень $x_m + y_n$, $m, n = \overline{1, \infty}$ з ймовірностями $P\{\xi = x_m, \eta = y_n\}$, які позначимо через p_{mn}. Далі маємо:

$$E(\xi + \eta) = \sum_{m,n=1}^{\infty} (x_m + y_n) p_{mn} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_m p_{mn} + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} y_n p_{mn} =$$

$$= \sum_{m=1}^{\infty} x_m p_m^{\xi} + \sum_{n=1}^{\infty} y_n p_n^{\eta} = E\xi + E\eta$$

З даної рівності, крім того видно, що із існування $E\xi$ та $E\eta$ випливає існування $E(\xi+\eta)$.

Якщо ж ξ та η – неперервні випадкові величини із щільностями розподілів $p_\xi(x)$ та $p_\eta(x)$ відповідно, то, як було показано, $p_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} p_{\xi,\eta}(u, x-u)du$, де $p_{\xi,\eta}(x,y)$ – сумісна щільність розподілу випадкових величин ξ та η. Отже, $E(\xi+\eta) = \int_{-\infty}^{+\infty} x p_{\xi+\eta}(x)dx = \int_{-\infty}^{+\infty} xdx \int_{-\infty}^{+\infty} p_{\xi,\eta}(u, x-u)du$. В останньому інтегралі зробимо заміну змінної x: $x-u=z$, тоді маємо:

$$E(\xi+\eta) = \int_{-\infty}^{+\infty}(u+z)dz \int_{-\infty}^{+\infty} p_{\xi,\eta}(u,z)du = \int_{-\infty}^{+\infty} udu \int_{-\infty}^{+\infty} p_{\xi,\eta}(u,z)dz + \int_{-\infty}^{+\infty} zdz \int_{-\infty}^{+\infty} p_{\xi,\eta}(u,z)du =$$

$$= \int_{-\infty}^{+\infty} up_\xi(u)du + \int_{-\infty}^{+\infty} zp_\eta(z)dz = E\xi + E\eta$$

5. Нехай ξ – дискретна випадкова величина з розподілом $P\{\xi = x_k\} = p_k, k = 1,2,\ldots$, для якої існує $E\xi$. Тоді, очевидно, що випадкова величина $g(\xi)$ набуває значень $g(x_k), k = 1,2,\ldots$ з тими самими ймовірностями: $P\{g(\xi) = g(x_k)\} = p_k, k = 1,2,\ldots$ Отже, $Eg(\xi) = \sum_{k=1}^{\infty} g(x_k)p_k$, що й треба було довести.

6. Для доведення властивості 6, скористаємось попередньою властивістю, вважаючи її вірною для будь-якої, а не лише дискретної випадкової величини. Припустимо, що ξ та η – неперервні випадкові величини із щільностями розподілів $p_\xi(x)$ та $p_\eta(x)$ відповідно. Представимо $E(\xi-\eta) = Eg(\xi,\eta)$, де $g(x,y) = x\cdot y$. Тоді $Eg(\xi,\eta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)p_{\xi,\eta}(x,y)dxdy$, а оскільки випадкові величини ξ та η – незалежні, то $p_{\xi,\eta}(x,y) = p_\xi(x)p_\eta(y)$, тому

$$Eg(\xi,\eta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp_\xi(x)p_\eta(y)dxdy = \int_{-\infty}^{+\infty} xp_\xi(x)dx \cdot \int_{-\infty}^{+\infty} yp_\eta(y)dy = E\xi\cdot E\eta$$

Наслідок 1. Якщо існує $E\xi$, то випадкова величина $\xi - E\xi$ – центрована.

Наслідок 2. Нехай $\xi_1, \xi_2, \ldots, \xi_n$ – інтегровані випадкові величини, $c_1, c_2, \ldots, c_n$ – довільні сталі, тоді існує математичне сподівання:

$$E(c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n) = \sum_{k=1}^{n} c_k E\xi_k .$$

Приклад 3. Нехай ξ – випадкова величина, що має біноміальний розподіл з параметрами n та p. Знайдемо її математичне сподівання. Для цього зауважимо, що випадкову величину ξ (яка означає кількість «успіхів» в серії незалежних випробувань Бернуллі) можна подати у вигляді суми n випадкових величин:

$\xi = \sum_{k=1}^{n} \xi_k$, де випадкова величина ξ_k набуває значення 1, якщо в k–му випробуванні стався «успіх», та значення 0, якщо сталась «невдача». Зрозуміло, що для будь-якого k: $E\xi_k = 1 \cdot p + 0 \cdot q = p$, а отже, $E\xi = \sum_{k=1}^{n} E\xi_k = np$

Приклад 4. Нехай ξ – випадкова величина, що має розподіл Пуассона з параметром λ. Знайдемо її математичне сподівання:

$$E\xi = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Приклад 5. Нехай випадкова величина ξ має нормальний розподіл $N(a, \sigma^2)$. Тоді

$$E\xi = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) dx = \left| \begin{matrix} x-a=y, \\ dx=dy \end{matrix} \right| = \int_{-\infty}^{+\infty} (a+y) \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{y^2}{2\sigma^2}\right) dy =$$

$$= a \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{y^2}{2\sigma^2}\right) dy + \int_{-\infty}^{+\infty} y \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{y^2}{2\sigma^2}\right) dy = a,$$

оскільки другий інтеграл рівний нулю, як інтеграл від непарної функції по симетричному проміжку інтегрування, а перший інтеграл рівний 1, як інтеграл від щільності нормального $N(0, \sigma^2)$ розподілу.

Таким чином, зрозумілий імовірнісний зміст параметру a нормального розподілу – це математичне сподівання даної випадкової величини.

Приклад 6. Нехай ξ – випадкова величина, що має розподіл Коші: $p_\xi(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}, x \in \mathbf{R}$. Знайти математичне сподівання випадкової величини $\eta = \min(1, |\xi|)$.

За властивістю 5,

$$E\eta = E\min(1, |\xi|) = \int_{-\infty}^{+\infty} \min(1, |x|) \cdot \frac{1}{\pi} \cdot \frac{1}{1+x^2} dx = \frac{2}{\pi} \int_{0}^{+\infty} \min(1, x) \cdot \frac{1}{1+x^2} dx =$$

$$= \frac{2}{\pi} \left(\int_0^1 \frac{x}{1+x^2} dx + \int_1^{+\infty} 1 \cdot \frac{x}{1+x^2} dx \right) = \frac{2}{\pi} \left(\frac{1}{2} \ln(1+x^2) \Big|_0^1 + \operatorname{arctg} x \Big|_1^{+\infty} \right) = \frac{\ln 2}{\pi} + \frac{1}{2}.$$

Зауваження 3. До речі, випадкова величина, розподілена за законом Коші, – приклад неінтегровної випадкової величини: $E\xi = \infty$.

2. Дисперсія випадкової величини та її властивості.

Означення 3. ***Дисперсією*** випадкової величини ξ називається число $D\xi$ (якщо воно існує), рівне:

$$D\xi = E(\xi - M\xi)^2. \qquad (6)$$

Зауваження 4. Дисперсію випадкової величини називають також ***другим центральним моментом***. (Центральним – тому що величина $\xi - M\xi$ центрована). Взагалі, якщо існує математичне сподівання $E(\xi^k)$, то воно зветься k-тим ***моментом*** випадкової величини.

Зауваження 5. Дисперсія випадкової величини вочевидь характеризує міру розсіяння її значень навколо її середнього.

Означення 4. Випадкова величина, дисперсія якої рівна одиниці, називається ***нормованою***.

Означення 5. Число $\sigma = \sqrt{D\xi}$ називається ***середньоквадратичним*** або ***стандартним відхиленням*** випадкової величини ξ.

Теорема 2. (Властивості дисперсії).

1. Для будь-якої випадкової величини, $D\xi \geq 0$.
2. $Dc = 0$
3. $D\xi = M\left(\xi^2\right) - \left(M\xi\right)^2$. (7)
4. $D\left(c\xi\right) = c^2 D\xi$. (8)
5. Якщо ξ та η – незалежні випадкові величини, які мають дисперсію, то існує $D\left(\xi + \eta\right) = D\xi + D\eta$,

Доведення. Властивості **1. – 4.** доведіть самостійно. Доведемо властивість **5.** За означенням,

$$D\left(\xi+\eta\right) = E\left[\left(\xi+\eta\right) - E\left(\xi+\eta\right)\right]^2 = E\left[\left(\xi - E\xi\right) + \left(\eta - E\eta\right)\right]^2 =$$
$$= E\left(\xi - E\xi\right)^2 + E\left(\eta - E\eta\right)^2 + 2E\left(\xi - E\xi\right)E\left(\eta - E\eta\right) =$$
$$= D\xi + D\eta + 2E\left[\left(\xi - E\xi\right)\cdot\left(\eta - E\eta\right)\right] \quad (9)$$

Якщо випадкові величини ξ та η незалежні, то і випадкові величини $\xi - E\xi$ та $\eta - E\eta$ – незалежні також, тому $E\left[\left(\xi - E\xi\right)\cdot\left(\eta - E\eta\right)\right] = E\left(\xi - E\xi\right)\cdot E\left(\eta - E\eta\right) = 0$, отже, ***для незалежних величин дисперсія суми рівна сумі дисперсій***: $D\left(\xi + \eta\right) = D\xi + D\eta$.

Приклад 7. Нехай ξ – випадкова величина, що має біноміальний розподіл з параметрами n та p. Відомо, що $\xi = \sum_{k=1}^{n} \xi_k$ (див. **Приклад 3**.), причому величини $\xi_k, k = \overline{1,n}$ – незалежні, як результати окремих незалежних дослідів. Тому $D\xi = \sum_{k=1}^{n} D\xi_k$. Далі, $D\xi_k = E\left(\xi_k - E\xi_k\right)^2 = E\left(\xi_k - p\right)^2 = E\left(\xi_k\right)^2 + p^2 - 2pE\xi_k =$
$= p + p^2 - 2p^2 = p - p^2 = pq$, отже, $D\xi = npq$

Приклад 8. Нехай випадкова величина ξ має нормальний розподіл $N\left(a,\sigma^2\right)$. Тоді, як відомо, $E\xi = a$, обчислимо тепер дисперсію ξ:

$$D\xi = M\left(\xi - a\right)^2 = \int_{-\infty}^{+\infty} \left(x-a\right)^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{\left(x-a\right)^2}{2\sigma^2}\right) dx = \left| \begin{array}{l} \dfrac{x-a}{\sigma\sqrt{2}} = \sqrt{y}, \\ dx = \dfrac{\sigma dy}{\sqrt{2y}} \end{array} \right| =$$

$$= 2\int_{0}^{+\infty} 2\sigma^2 y \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-y} \cdot \frac{\sigma dy}{\sqrt{2y}} =$$

$$= 2\sigma^2 \cdot \frac{1}{\sqrt{\pi}} \int_{0}^{+\infty} y^{1/2} e^{-y} dy = 2\sigma^2 \cdot \frac{1}{\sqrt{\pi}} \cdot \Gamma\left(\frac{3}{2}\right) = \sigma^2 \cdot \frac{1}{\sqrt{\pi}} \cdot \Gamma\left(\frac{1}{2}\right) = \sigma^2$$

Таким чином, параметр σ^2 нормального розподілу – це дисперсія даної випадкової величини.

Приклад 9. Нехай випадкова величина ξ має середнє $E\xi = a$ та дисперсію $D\xi = \sigma^2$. Неважко зрозуміти тоді, що випадкова величина $\eta = \dfrac{\xi - a}{\sigma}$ – центрована та нормована, оскільки

$$E\eta = E\frac{\xi - a}{\sigma} = \frac{1}{\sigma}E(\xi - a) = 0,\ D\eta = D\frac{\xi - a}{\sigma} = \frac{1}{\sigma^2}D(\xi - a) = \frac{D\xi}{\sigma^2} = 1.$$

3. Коваріація та коефіцієнт кореляції двох випадкових величин.

Означення 6. ***Коваріацією*** випадкових величин ξ та η називається число $\operatorname{cov}(\xi,\eta)$ (якщо воно існує), рівне $\operatorname{cov}(\xi,\eta) = E[(\xi - E\xi)\cdot(\eta - E\eta)]$. (10)

Неважко перевірити, що $\operatorname{cov}(\xi,\eta) = E(\xi\cdot\eta) - E\xi\cdot E\eta$. (11)

Зауваження 6. Вище було доведено, що для двох довільних випадкових величин $D(\xi+\eta) = D\xi + D\eta + 2\operatorname{cov}(\xi,\eta)$ (див. формулу (9)).

Зауваження 7. Також було доведено, що, якщо ξ та η – незалежні випадкові величини, то $\operatorname{cov}(\xi,\eta) = 0$.

Означення 7. ***Коефіцієнтом кореляції*** випадкових величин ξ та η з ненульовими дисперсіями називається число $\mathrm{r}(\xi,\eta)$ (якщо воно існує), рівне

$$\mathrm{r}(\xi,\eta) = \frac{\operatorname{cov}(\xi,\eta)}{\sqrt{D\xi}\ \sqrt{D\eta}}. \quad (12)$$

Означення 8. Випадкові величини ξ та η називаються ***некорельованими***, якщо існує $\mathrm{r}(\xi,\eta) = 0$.

В наступній теоремі зібрані властивості коефіцієнта кореляції двох випадкових величин. В усіх твердженнях припускається існування $\mathrm{r}(\xi,\eta)$.

Теорема 3. (Властивості коефіцієнта кореляції).

1. Для будь-яких випадкових величин $|\mathrm{r}(\xi,\eta)| \le 1$.
2. Якщо ξ та η – ***незалежні*** випадкові величини, то вони ***некорельовані.*** Обернене твердження, взагалі кажучи, невірне.
3. Якщо (ξ,η) – випадковий вектор з нормальним розподілом з параметрами $a_1, a_2, \sigma_1^2, \sigma_2^2$ та r (див. попередню лекцію), то а) параметр $r = \mathrm{r}(\xi,\eta)$; б) з некорельованості компонентів, тобто $r = 0$, випливає, що ξ та η – незалежні випадкові величини. Таким чином, для нормально розподілених випадкових величин незалежність та некорельованість – еквівалентні.

Доведення. Твердження 2 вже обговорене у зауваженні 7. Твердження 3 перевірте самостійно. Доведемо властивість 1. З цією метою спрощуватимемо очевидну нерівність:

$$E\left[\frac{(\xi - E\xi)}{\sqrt{D\xi}} \pm \frac{(\eta - E\eta)}{\sqrt{D\eta}}\right]^2 \ge 0 \quad (13)$$

$$\Rightarrow\ E\frac{(\xi - E\xi)^2}{D\xi} + E\frac{(\eta - E\eta)^2}{D\eta} \pm 2E\frac{(\xi - E\xi)}{\sqrt{D\xi}}\cdot\frac{(\eta - E\eta)}{\sqrt{D\eta}} \ge 0 \Rightarrow\ 2 \pm 2\,\mathrm{r}(\xi,\eta) \ge 0,$$ звідки й випливає потрібна нерівність.

Зауваження 8. З нерівності (13) додатково випливає, що $|\mathrm{r}(\xi,\eta)| = 1$ тоді і тільки тоді, коли одна з випадкових величин лінійно виражається через іншу, наприклад,

$\xi = A\eta + B$, де A та B – деякі числа (причому $\operatorname{sign} A = \operatorname{sign}(\mathrm{r}(\xi,\eta))$). Тобто коефіцієнт кореляції рівний ± 1 вказує на лінійну залежність випадкових величин ξ та η. Проте, наступний приклад демонструє, що випадкові величини, пов'язані ***нелінійною залежністю*** можуть бути некорельовані, тобто коефіцієнт кореляції проявляє чутливість саме до лінійної залежності випадкових величин.

Приклад 10. Нехай випадкова величина ξ – рівномірно розподілена на відрізку $[0,2\pi]$. Розглянемо випадкові величини $\eta_1 = \sin\xi$ та $\eta_2 = \cos\xi$. Очевидно, що $\eta_1{}^2 + \eta_2{}^2 = 1$, тобто величини η_1, η_2 – функціонально залежні.

Обчислимо $E\eta_1 = \int\limits_0^{2\pi} \frac{\sin x}{2\pi} dx = 0$. Аналогічно, $E\eta_2 = 0$. Крім того, $E\eta_1 \cdot \eta_2 = E\sin\xi \cdot \cos\xi = \frac{1}{2} E\sin 2\xi = \frac{1}{2}\int\limits_0^{2\pi} \frac{\sin 2x}{2\pi} dx = 0$, тобто з формул (11) та (12) випливає, що $\mathrm{r}(\eta_1, \eta_2) = 0$ – величини η_1, η_2 є некорельованими, проте залежними.

Лекція 7. Нерівність Чебишова. Закон великих чисел та його наслідки. Характеристична функція та її властивості.

Одним з найбільш переконливих проявів закономірностей, що виникають при спостереженні великої кількості випадкових величин, є теореми про закони «великих чисел». Ці теореми стверджують, що конкретні особливості окремих випадкових явищ майже не впливають на їх ***середній результат***, якщо кількість цих явищ досить велика. В цій стійкості середніх і полягає фізичний зміст законів великих чисел. Зокрема, нам відомий той факт, що в схемі незалежних випробувань Бернуллі частота успіху, яка є випадковою величиною, приймає значення, близькі до сталої ймовірності успіху p, якщо кількість випробувань досить велика.

1. Нерівність Чебишова.

Доведення законів великих чисел базується на нерівності, вперше надрукованій російською та французькою мовами у 1867 році в роботі Пафнутія Львовича Чебишова «Про середні величини».

Твердження 1. (Нерівність Чебишова).
Для довільної випадкової величини ξ із скінченою дисперсією $D\xi$ при будь-якому значенні $\varepsilon > 0$ виконується нерівність

$$P\{|\xi - E\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2} \qquad (1)$$

Доведення. Розглянемо випадок, коли ξ – неперервна випадкова величина із щільністю розподілу $p(x)$. Позначимо $E\xi = a$. Тоді

$$P\{|\xi - E\xi| \geq \varepsilon\} = \int\limits_{\{x:|x-a|\geq\varepsilon\}} p(x)dx \leq \int\limits_{\{x:|x-a|\geq\varepsilon\}} \frac{|x-a|^2}{\varepsilon^2} p(x)dx \leq \frac{1}{\varepsilon^2}\int\limits_{-\infty}^{+\infty} |x-a|^2 p(x)dx = \frac{D\xi}{\varepsilon^2}.$$

У випадку дискретної випадкової величини доведення цілком аналогічне із заміною інтегралів на відповідні суми – спробуйте провести його самостійно.

Приклад 1. Нехай ξ – випадкова величина із математичним сподіванням $E\xi = a$ та дисперсією $D\xi = \sigma^2$. З нерівності Чебишова випливає, що

$P\{|\xi - a| \geq 3\sigma\} \leq \frac{D\xi}{9\sigma^2} = \frac{1}{9}$. Якщо припустити, що ξ має нормальний розподіл, то значення лівої частини нерівності можна визначити по таблицях: $P\{|\xi - a| \geq 3\sigma\} \approx 0{,}0027$, в той час, як значення оцінки нерівності Чебишова «гірше» майже у 40 разів. Проте, значимість даної нерівності саме в її універсальності та незалежності від характеру розподілу випадкової величини.

2. Закон великих чисел.

Теорема 1. (Теорема Чебишова про закон великих чисел).
Нехай $\{\xi_n\}, n \geq 1$ – послідовність незалежних випадкових величин, що мають дисперсії, обмежені сталою величиною: $\forall n \geq 1 \ D\xi_n \leq c$. Тоді

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}E\xi_k\right| < \varepsilon\right\} = 1 \qquad (2)$$

Доведення. Розглянемо випадкову величину $\eta_n = \frac{\sum_{k=1}^{n}\xi_k}{n}$. Її середнє рівне: $E\eta_n = \frac{\sum_{k=1}^{n}E\xi_k}{n}$, а дисперсія – $D\eta_n = \frac{D\left(\sum_{k=1}^{n}\xi_k\right)}{n^2} = \frac{\sum_{k=1}^{n}D\xi_k}{n^2} \leq \frac{cn}{n^2} = \frac{c}{n}$. Застосуємо до величини η_n нерівність Чебишова:

$P\{|\eta_n - E\eta_n| < \varepsilon\} = 1 - P\{|\eta_n - E\eta_n| \geq \varepsilon\} \geq 1 - \frac{D\eta_n}{\varepsilon^2} \geq 1 - \frac{c}{n\varepsilon^2}$. Перейдемо до границі при $n \to \infty$: $\lim_{n\to\infty} P\{|\eta_n - M\eta_n| < \varepsilon\} \geq 1$, звідки й випливає, що

$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}E\xi_k\right| < \varepsilon\right\} = 1.$

Зауваження. Теорема Чебишова стверджує, таким чином, що середнє арифметичне великої кількості незалежних випадкових величин проявляє статистичну стійкість – поводить себе майже як невипадкова стала величина рівна середньому арифметичному математичних сподівань випадкових величин.

Розглянемо для прикладу певний об'єм газу, що міститься у деякому сосуді. Його молекули перебувають у постійному хаотичному русі, причому для кожної окремої молекули неможливо передбачити її місцезнаходження та швидкість у деякий момент часу. Проте можна оцінити долю частинок газу, що рухаються з деякою заданою швидкістю або містяться у певному об'ємі – саме це і цікавить фізика. Якщо ж говорити про тиск цього газу, то він визначається середньою кількістю ударів молекул різних швидкостей, що приходяться на пластинку одиничної площі за одиницю часу. Зрозуміло, що тут ми маємо справу з випадковими величинами, проте, як відомо, тиск газу – величина стала у незмінних умовах. Це прояв усереднення, яке є наслідком закону великих чисел. Саме фізичні явища яскраво демонструють дію закону великих чисел.

Означення 1. Будемо говорити, що послідовність випадкових величин $\{\xi_n\}, n \geq 1$ ***збігається за ймовірністю*** до випадкової величини ξ при $n \to \infty$, позначаючи цей факт наступним чином: $\xi_n \xrightarrow{p} \xi,\ n \to \infty$, якщо $\lim\limits_{n\to\infty} P\{|\xi_n - \xi| < \varepsilon\} = 1\ \forall \varepsilon > 0$.

Зауваження 1. Збіжність за ймовірністю $\xi_n \xrightarrow{p} \xi,\ n \to \infty$ ***не означає***, що значення випадкових величин прямують до значень випадкової величини ξ при $n \to \infty$, а означає лише, що відхилення цих значень більші як завгодно малої величини ε мають нульову ймовірність – проте з'явлення такого відхилення не є неможливою подією.

Зауваження 2. Теорема Чебишова стверджує, таким чином, що

$$\frac{\sum\limits_{k=1}^{n} \xi_k}{n} \xrightarrow{p} \frac{\sum\limits_{k=1}^{n} E\xi_k}{n},\ n \to \infty.$$

Зауваження 3. Теорема Чебишова для ***незалежних однаково розподілених*** випадкових величин $\{\xi_n\}, n \geq 1$ з $E\xi_n = a, D\xi_n = \sigma^2$ стверджує, що

$$\frac{\sum\limits_{k=1}^{n} \xi_k}{n} \xrightarrow{p} a,\ n \to \infty. \qquad (3)$$

Зауваження 4. Вже відома нам теорема Бернуллі таким чином, є прямим наслідком теореми Чебишева про закон великих чисел. Сформулюємо її тут ще раз.

Теорема 2. (Теорема Бернуллі про закон великих чисел).
Нехай $\{\nu_n\}, n \geq 1$ – кількість успіхів в серії з n незалежних випробувань Бернуллі із сталою ймовірністю успіху окремого випробування рівною p. Тоді $\lim\limits_{n\to\infty} P\left\{\left|\frac{\nu_n}{n} - p\right| < \varepsilon\right\} = 1$, тобто $\frac{\nu_n}{n} \xrightarrow{p} p,\ n \to \infty$ (4)

Доведення. Розглянемо випадкові величини ξ_k, $1 \leq k \leq n$, які набувають значення 1, якщо в k–му випробуванні стався «успіх», та значення 0, якщо сталась «невдача». Було визначено (див. попередню лекцію), що $E\xi_k = p$, а $\nu_n = \sum\limits_{k=1}^{n} \xi_k$. Тоді твердження (4) прямо випливає з (3).

Наслідок 1. (Застосування закону великих чисел для вимірів фізичних величин).
Нехай ν – деяка фізична величина з теоретичним значенням a, яке необхідно визначити шляхом фізичного експерименту. Останній полягає у проведенні n незалежних вимірів цієї величини, наслідком яких є випадкові величини $\{\nu_n\}, n \geq 1$. Зрозуміло, що ці випадкові величини – незалежні та однаково розподілені з величиною ν. Причому $E\nu_n = a$ внаслідок відсутності систематичної похибки вимірювання, а $D\nu_n = D\nu = \sigma^2$. Розглянемо випадкову величину $\eta_n = \frac{\sum\limits_{k=1}^{n} \nu_k}{n}$. Для

неї $E\eta_n = a$, $D\eta_n = \frac{1}{n^2}\sum_{k=1}^{n} D\nu_k = \frac{\sigma^2}{n}$. Отже, застосувавши теоремиу Чебишова одержуємо, що $\frac{\sum_{k=1}^{n}\nu_k}{n} \xrightarrow{p} a,\ n\to\infty$, тому шукане значення a величини ν може бути досить точно замінене середнім арифметичним послідовності вимірів цієї величини $\{\nu_n\}, n\ge 1$ для великих значень n.

Наслідок 2. (Метод Монте-Карло обчислення інтегралів виду $\int_0^1 f(x)dx$).

Нехай $\{\xi_n\}, n\ge 1$ – послідовність незалежних випадкових величин рівномірно розподілених на відрізку [0,1]. Якщо $f(x)$ – деяка борелівська функція, то $\{f(\xi_n)\}, n\ge 1$ – також незалежні випадкові величини з $Ef(\xi_n) = \int_0^1 f(x)dx$. Отже, $\sum_{k=1}^{n} f(\xi_k) \Big/ n \xrightarrow{p} \int_0^1 f(x)dx,\ n\to\infty$. Таким чином, для обчислення значення інтегралу виду $\int_0^1 f(x)dx$ можна запропонувати наступний метод: змоделювати досить велику кількість n рівномірно розподілених на інтервалі [0,1] випадкових величин $\{\xi_k\}, 1\le k\le n$ та обчислити значення інтеграла як наближення $\int_0^1 f(x)\,dx \approx \frac{\sum_{k=1}^{n} f(\xi_k)}{n}$.

3. Характеристична функція випадкової величини та її властивості.

Нехай ξ – випадкова величина на даному імовірнісному просторі $(\Omega, \Im, P)$.

Означення 2. ***Характеристичною функцією*** випадкової величини ξ називається комплекснозначна функція дійсного аргументу

$$\varphi_\xi(t) = Ee^{it\xi},\quad t\in\mathbf{R} \tag{5}$$

Зауважимо, що характеристична функція визначена для будь-якої випадкової величини в силу нерівності $\left|e^{it\xi}\right| \le 1\ \forall t\in\mathbf{R}$. Якщо згадати позначення, введені у лекції 6, можна написати, що

$$\varphi_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} dF_\xi(x) \tag{6}$$

Зауваження 5. З рівності (6) випливає, що характеристична функція випадкової величини є перетворенням Фур'є її розподілу.

Приклад 2. Нехай ξ – випадкова величина рівномірно розподілена на симетричному проміжку $[-a, a]$. Тоді

$\varphi_\xi(t)=\int\limits_{-\infty}^{+\infty}e^{itx}p_\xi(x)dx=\frac{1}{2a}\int\limits_{-a}^{+a}e^{itx}dx=\frac{e^{ita}-e^{-ita}}{2a\cdot it}=\frac{\sin at}{at}$ – характеристична функція цієї випадкової величини має дійсні значення.

Приклад 3. Нехай ξ – випадкова величина розподілена за законом Пуассона з параметром λ. Знайдемо її характеристичну функцію:

$$\varphi_\xi(t)=Ee^{it\xi}=\sum_{n=0}^{\infty}e^{itn}\frac{e^{-\lambda}\lambda^n}{n!}=e^{-\lambda}\sum_{n=0}^{\infty}\frac{\left(e^{it}\lambda\right)^n}{n!}=e^{-\lambda}\cdot e^{\lambda e^{it}}=\exp\{\lambda\left(e^{it}-1\right)\}$$

Теорема 3. (Властивості характеристичної функції).

1. $\left|\varphi_\xi(t)\right|\le 1 \quad \forall t\in\mathbf{R}; \quad \varphi_\xi(0)=1$
2. Якщо ξ та η – незалежні випадкові величини, то $\varphi_{\xi+\eta}(t)=\varphi_\xi(t)\cdot\varphi_\eta(t)$.
3. $\varphi_{a\xi+b}(t)=e^{ibt}\varphi_\xi(at)$ для довільної випадкової величини ξ та будь-яких констант a і b.
4. $\varphi_\xi(t)$ – рівномірно неперервна функція на $\mathbf{R}$, тобто $\forall\varepsilon>0\;\exists\delta=\delta(\varepsilon):\forall t_1,t_2:|t_1-t_2|<\delta\Rightarrow\left|\varphi_\xi(t_1)-\varphi_\xi(t_2)\right|<\varepsilon$.
5. Нехай випадкова величина ξ має k-ий момент: $\exists m_k=E\xi^k$. Тоді існує k-та похідна характеристичної функції і для всіх $n\le k$ виконується рівність $\varphi_\xi{}^{(n)}(0)=i^nE\xi^n$, тобто $E\xi^n=\frac{\varphi_\xi{}^{(n)}(0)}{i^n}\forall n\le k$.
6. $\varphi_\xi(-t)=\overline{\varphi_\xi(t)}$
7. $\varphi_\xi(t)$ – ***додатно визначена*** функція на $\mathbf{R}$, тобто $\forall n\in\mathbf{N}\;\forall t_1,t_2,\ldots,t_n\in\mathbf{R}\;\forall c_1,c_2,\ldots,c_n\in\mathbf{C}:\sum\limits_{k=1}^{n}\sum\limits_{j=1}^{n}\varphi_\xi\left(t_k-t_j\right)c_k\overline{c_j}\ge 0$

Доведення. Властивість перша очевидно випливає з означення характеристичної функції.

2. Розглянемо для незалежних випадкових величин ξ та η характеристичну функцію їх суми: $\varphi_{\xi+\eta}(t)=Ee^{it(\xi+\eta)}=Ee^{it\xi}\cdot e^{it\eta}=Ee^{it\xi}\cdot Ee^{it\eta}=\varphi_\xi(t)\cdot\varphi_\eta(t)$.
3. Розглянемо $\varphi_{a\xi+b}(t)=Ee^{it(a\xi+b)}=Ee^{i(ta)\xi}\cdot e^{itb}=e^{itb}\cdot\varphi_\xi(at)$.
4. Нехай $\varepsilon>0$ задано. Позначимо для довільних $t_1<t_2$: $h=t_2-t_1$. Одержимо $\varphi_\xi(t_2)-\varphi_\xi(t_1)=\int\limits_{-\infty}^{+\infty}\left(e^{it_2x}-e^{it_1x}\right)dF_\xi(x)=\int\limits_{-\infty}^{+\infty}e^{it_1x}\cdot\left(e^{ihx}-1\right)dF_\xi(x)$. Оскільки $\left|e^{ihx}-1\right|\le 2$, а $\int\limits_{-\infty}^{+\infty}dF_\xi(x)=1$, знайдеться така величина $A>0$, що $\int\limits_{|x|>A}dF_\xi(x)<\varepsilon/4$. Далі, для всіх $x:|x|<A$ знайдемо таке $h>0$, що

$\left|e^{ihx}-1\right| \le \varepsilon/2$. Це можливо, оскільки функція e^{ihx} – рівномірно неперервна на $\mathbf{R}$. Тоді $\left|\varphi_\xi(t_2)-\varphi_\xi(t_1)\right| \le$

$$\le \int_{-\infty}^{+\infty}\left|e^{ihx}-1\right|dF_\xi(x) < 2\cdot\int_{|x|>A} dF_\xi(x)+\int_{-A}^{+A}\varepsilon/2\cdot dF_\xi(x) < 2\cdot\varepsilon/4+\varepsilon/2 = \varepsilon.$$

Таким чином, для заданого $\varepsilon>0$ можна визначити таке число $\delta = h$, що з нерівності $|t_1-t_2|<\delta$ випливає нерівність $\left|\varphi_\xi(t_1)-\varphi_\xi(t_2)\right|<\varepsilon$.

5. Припустимо, що існує $m_k = E\xi^k$ при деякому значенні k. Це означає абсолютну збіжність інтегралу : $\int_{-\infty}^{+\infty}\left|x^k\right|dF_\xi(x)<\infty$. Розглянемо рівність $\varphi_\xi(t)=\int_{-\infty}^{+\infty}e^{itx}dF_\xi(x)$ та візьмемо n-ту похідну (де $n\le k$) від цієї рівності (правомірність цих дій буде обґрунтовано нижче):

$$\varphi_\xi^{(k)}(t)\big|_{t=0} = \int_{-\infty}^{+\infty}(ix)^k e^{itx}dF_\xi(x)\big|_{t=0} = i^k\int_{-\infty}^{+\infty}x^k e^{itx}dF_\xi(x)\big|_{t=0} = i^k E\xi^k = i^k m_k.$$

Оскільки $\left|\int_{-\infty}^{+\infty}x^k e^{itx}dF_\xi(x)\right| \le \int_{-\infty}^{+\infty}\left|x^k\right|dF_\xi(x)<\infty$, то інтеграл $\int_{-\infty}^{+\infty}x^k e^{itx}dF_\xi(x)$ від параметра t рівномірно збіжний за ознакою Вейерштрасса, тому диференціювання було законним.

6. Дійсно, $\varphi_\xi(-t)=\int_{-\infty}^{+\infty}e^{-itx}dF_\xi(x)=\int_{-\infty}^{+\infty}\overline{e^{itx}}dF_\xi(x)=\overline{\int_{-\infty}^{+\infty}e^{itx}dF_\xi(x)}=\overline{\varphi_\xi(t)}$

7. Розглянемо

$$\sum_{k=1}^{n}\sum_{j=1}^{n}\varphi_\xi(t_k-t_j)c_k\overline{c_j} = \sum_{k=1}^{n}\sum_{j=1}^{n}Ee^{i(t_k-t_j)\xi}\,c_k\overline{c_j} = E\left(\sum_{k=1}^{n}e^{it_k\xi}\,c_k\cdot\sum_{j=1}^{n}e^{-it_j\xi}\,\overline{c_j}\right) =$$

$$= E\left(\sum_{k=1}^{n}e^{it_k\xi}\,c_k\cdot\overline{\sum_{j=1}^{n}e^{it_j\xi}\,c_j}\right) = E\left|\sum_{k=1}^{n}e^{it_k\xi}\,c_k\right|^2 \ge 0$$

Приклад 3 (продовження). Нехай ξ – випадкова величина розподілена за законом Пуассона з параметром λ. Була визначена її характеристична функція $\varphi_\xi(t)=\exp\{\lambda(e^{it}-1)\}$. Згідно властивості **5.**, якщо існує математичне сподівання величини ξ, то його знаходимо за формулою

$$E\xi = \frac{\varphi_\xi'(0)}{i} = -i\exp\{\lambda(e^{it}-1)\}\lambda e^{it}i\big|_{t=0} = \lambda.$$

Приклад 4. Нехай ξ – випадкова величина, що має нормальний розподіл $N(a,\sigma^2)$. Визначимо її характеристичну функцію. Зауважимо спочатку, що випадкова величина $\eta = \frac{\xi-a}{\sigma}$ має стандартний нормальний розподіл $N(0,1)$. Її

характеристична функція: $\varphi_\eta(t) = Me^{it\eta} = \int\limits_{-\infty}^{+\infty} e^{itx} p_\eta(x)dx = \int\limits_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx.$

Оскільки існує $D\eta$, функція $\varphi_\eta(t)$ має і першу , і другу похідні, тоді

$$\varphi_\eta{}'(t) = \int\limits_{-\infty}^{+\infty} ixe^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = -\frac{i}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{itx} d\left(e^{-x^2/2} \right) =$$

$$= -\frac{i}{\sqrt{2\pi}} e^{itx} e^{-x^2/2} \Big|_{x\to-\infty}^{x\to+\infty} + \frac{i}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-x^2/2} ite^{itx} dx = -t \int\limits_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \quad ,$$ тобто функція $\varphi_\eta(t)$ задовольняє диференціальне рівняння: $\frac{d\varphi_\eta(t)}{dt} = -t\varphi_\eta(t)$. Його розв'язками є сімейство функцій $\varphi_\eta(t) = ce^{-t^2/2}$. Оскільки $\varphi_\eta(0) = 1$, то $c = 1$. Отже, якщо випадкова величина η має стандартний нормальний розподіл, то її характеристична функція рівна $\varphi(t) = e^{-t^2/2}$. Розглянемо тепер випадкову величину $\xi = \sigma\eta + a$ з розподілом $N(a, \sigma^2)$. Згідно властивості **3.**,

$$\varphi_\xi(t) = e^{iat} \cdot e^{-\sigma^2 t^2/2} = \exp\left\{ iat - \frac{\sigma^2 t^2}{2} \right\} \qquad (7)$$

Позначимо $\phi(t) = \ln \varphi_\xi(t)$. Тоді з припущення про існування моментів відповідних порядків випливає, що існує похідна $\phi'(t) = \left[\ln \varphi_\xi(t)\right]' = \frac{\varphi'_\xi(t)}{\varphi_\xi(t)}$. Отже, $\phi'(0) = \frac{\varphi'_\xi(0)}{\varphi_\xi(0)} = iE\xi$. Аналогічно, $\phi''(t) = \left[\ln \varphi_\xi(t)\right]'' = \frac{\varphi''_\xi(t)\varphi_\xi(t) - \left[\varphi'_\xi(t)\right]^2}{\varphi_\xi(t)}$, тому $\phi''(0) = i^2 E\xi^2 - \left[iE\xi\right]^2 = -D\xi$. Взагалі, виявляється, що числа $s_k = i^k \phi^{(k)}(0), k \in \mathbf{N}$ – дійсні. Вони називаються ***семі-інваріантами*** (***кумулянтами***) випадкової величини ξ. Вони є коефіцієнтами розкладу $\phi(t) = \ln \varphi_\xi(t)$ у степеневий ряд.

Сформулюємо (без доведення) дуже важливу теорему, яка пояснює етимологію терміну «характеристична функція».

Теорема 4. (єдиності для характеристичної функції).
Кожній характеристичній функції $\varphi_\xi(t)$ відповідає одна і тільки одна функція розподілу $F_\xi(x)$, тобто з того, що $\varphi_\xi(t) = \int\limits_{-\infty}^{+\infty} e^{itx} dF_1(x)$ і $\varphi_\xi(t) = \int\limits_{-\infty}^{+\infty} e^{itx} dF_2(x)$ випливає тотожність розподілів: $F_1(x) \equiv F_2(x)$.

Приклад 5. Нехай випадкові величини ξ та η мають нормальні розподіли $N(a_1, {\sigma_1}^2)$ та $N(a_2, {\sigma_2}^2)$. Тоді характеристичною функцією їх суми буде

$$\varphi_{\xi+\eta}(t) = \exp\left\{ ia_1 t - \frac{{\sigma_1}^2 t^2}{2} \right\} \cdot \exp\left\{ ia_1 t - \frac{{\sigma_1}^2 t^2}{2} \right\} = \exp\left\{ i(a_1 + a_2)t - \frac{\left({\sigma_1}^2 + {\sigma_2}^2\right)t^2}{2} \right\},$$

тобто сумарний розподіл також нормальний – $N\left(a_1+a_2,\ {\sigma_1}^2+{\sigma_2}^2\right)$ в силу теореми єдиності.

Означення 3. Нехай ξ – дискретна випадкова величина, що набуває цілих невід'ємних значень: $P\{\xi=k\}=p_k,\ k=0,1,2,\ldots$ ***Твірною (генераторною) функцією*** випадкової величини ξ називається комплексна функція

$$\psi_\xi(z)=Ez^\xi=\sum_{k=0}^{\infty}z^k p_k,\ z\in\mathbf{C},\ |z|\le 1. \tag{8}$$

Неважко зрозуміти, що $\psi_\xi\left(e^{it}\right)=\varphi_\xi(t)$. Твірна функція однозначно визначає розподіл: $p_k=\frac{1}{k!}\psi_\xi^{(k)}(0), k\ge 0$. Крім того, твірна функція суми незалежних випадкових величин ξ та η рівна добутку їх твірних: $\psi_{\xi+\eta}(z)=\psi_\xi(z)\cdot\psi_\eta(z)$.

Приклад 5. Нехай ξ – випадкова величина, що має розподіл Бернуллі з параметрами n та p. Визначимо її твірну функцію:

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

Приклад 6. Нехай ξ – випадкова величина, що має розподіл Пуассона з параметром λ. Її твірна функція є такою:

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

<u>Лекція 8</u>. Центральна гранична теорема. Поняття про випадковий процес.

Сформульована нами раніше інтегральна гранична теорема Муавра-Лапласа стверджувала, що із зростанням кількості випробувань n в схемі Бернуллі центрована та нормована сума випадкових величин $\sum_{k=1}^{n}\xi_k$, яка визначає кількість успіхів в серії n випробувань, поводить себе майже як випадкова величина із стандартним нормальним розподілом. Дійсно, враховуючи відомі значення числових характеристик біноміального розподілу можна записати:

$$P\left\{n_1\le\sum_{k=1}^{n}\xi_k\le n_2\right\}=P\left\{a\le\frac{\sum_{k=1}^{n}\xi_k-np}{\sqrt{npq}}\le b\right\}=P\left\{a\le\frac{\sum_{k=1}^{n}\xi_k-\sum_{k=1}^{n}E\xi_k}{\sqrt{\sum_{k=1}^{n}D\xi_k}}\le b\right\}\approx$$

$$\approx\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx, \tag{1}$$

де $a=\frac{n_1-np}{\sqrt{npq}},\ b=\frac{n_2-np}{\sqrt{npq}}$

Цілий ряд вчених, зокрема П.Л. Чебишов та його учні А.А. Марков і А.М. Ляпунов присвятили свої зусилля дослідженню того, наскільки формула (1) залежить від конкретного характеру розподілів величин ξ_k, та які вимоги необхідно накладати на ці випадкові величини, щоб було справедливим вказане наближення. Їх дослідження привели до ***центральної граничної теореми*** теорії ймовірностей, найпростіший варіант якої буде розглянуто нижче.

Грандіозність значення цієї теореми визначається тим фактом, що численні результати природничих та технічних наук підтверджують дію цього центрального закону теорії ймовірностей. Адже будь-який природний чи технічний процес відбувається під впливом великої кількості випадкових, діючих незалежно, факторів, кожний з яких здатний лише несуттєво змінити хід процесу. Тому спостерігач має справу із сумарною дією всіх цих факторів. Центральна гранична теорема дозволяє робити висновки про граничну поведінку всієї суми випадкових впливів. Виявляється, що граничний розподіл близький до стандартного нормального розподілу. (Саме цьому факту завдячує назва «стандартний нормальний розподіл»).

Доведення центральної граничної теореми базується на тому факті, що між функціями розподілу та характеристичними функціями існує взаємно однозначний зв'язок, який поширюється навіть на відповідні граничні функції при збіжності розподілів та характеристичних функцій. Наведемо спочатку декілька важливих для подальшого викладення охначень.

Означення 1. Послідовність функцій розподілу $F_n(x), n \geq 1$ випадкових величин $\{\xi_n\}, n \geq 1$ ***слабко збігається*** при $n \to \infty$ до функції розподілу $F(x)$ деякої випадкової величини ξ (цей вид збіжності позначається наступним чином $F_n(x) \underset{n\to\infty}{\overset{сл.}{\Rightarrow}} F(x)$), якщо для кожної обмеженої неперервної на **R** функції $f(x)$ виконується гранична рівність: $\lim\limits_{n\to\infty} \int\limits_{-\infty}^{+\infty} f(x) dF_n(x) = \int\limits_{-\infty}^{+\infty} f(x) dF(x)$.

Зауваження. $F_n(x) \underset{n\to\infty}{\overset{сл.}{\Rightarrow}} F(x)$ таким чином означає збіжність числової послідовності $Mf(\xi_n) \underset{n\to\infty}{\to} M\xi$.

Теорема 1. **(Леві).** Для того, щоб послідовність функцій розподілу $F_n(x)$ слабко збігалась до деякого розподілу $F_0(x)$, необхідно та достатньо, щоб послідовність характеристичних функцій $\varphi_n(t)$ розподілів $F_n(x)$ збігалась при кожному t до деякої функції $\varphi(t)$, неперервної в нулі. При цьому $\varphi(t)$ буде характеристичною функцією розподілу $F_0(x)$. (Цю теорему використовуємо без доведення)

Таким чином, слабка збіжність послідовності розподілів еквівалентна збіжності відповідних характеристичних функцій.

Теорема 2. **(Центральна гранична теорема для послідовності незалежних однаково розподілених випадкових величин).** Нехай $\{\xi_n\}, n \geq 1$ – послідовність незалежних випадкових величин, що мають однаковий

розподіл. Позначимо : $E\xi_n = a,\ D\xi_n = \sigma^2,\ \nu_n = \sum_{k=1}^{n} \xi_k \ \ \forall n = 1,2,\ldots$. Тоді

$$\lim_{n\to\infty} P\left\{\frac{\nu_n - E\nu_n}{\sqrt{D\nu_n}} < x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad \text{рівномірно по } x \in \mathbf{R}, \tag{2}$$

тобто послідовність розподілів $F_n(x)$ випадкових величин $\eta_n = \frac{\nu_n - E\nu_n}{\sqrt{D\nu_n}}$ слабко збігається до стандартного нормального розподілу $N(0,1)$.

Доведення. Позначимо через $\varphi_n(t)$ характеристичні функції випадкових величин η_n. Згідно теореми Леві, досить показати, що $\lim_{n\to\infty} \varphi_n(t) = e^{-\frac{t^2}{2}}$. Оскільки $\eta_n = \frac{\nu_n - E\nu_n}{\sqrt{D\nu_n}} = \sum_{k=1}^{n} \frac{\xi_k - a}{\sigma\sqrt{n}}$, то на підставі властивостей **2.** та **3.** характеристичної функції можна стверджувати, що $\varphi_n(t) = \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$, де $\varphi(t)$ – характеристична функція випадкової величини $\xi_n - a,\ n \ge 1$. Оскільки існує дисперсія $D\xi_n$, то існує друга похідна функції $\varphi(t)$. Скористаємось розкладом в ряд Тейлора при $n \to \infty$ (тоді $\frac{t}{\sigma\sqrt{n}} \to 0$) функції $\varphi(t)$ із залишковим членом у формі Пеано:

$$\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) = \varphi(0) + \varphi'(0)\frac{t}{\sigma\sqrt{n}} + \frac{\varphi''(0)}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + o\left(\frac{t}{\sigma\sqrt{n}}\right)^2 = 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{2n}\right)$$

Тут враховано, що $\varphi(0) = 1,\ \varphi'(0) = iE(\xi_k - a) = 0,\ \varphi''(0) = i^2 E(\xi_k - a)^2 = -\sigma^2$. Звідси випливає, що $\varphi_n(t) = \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{2n}\right)\right]^n \underset{n\to\infty}{\to} e^{-\frac{t^2}{2}}$ рівномірно по $t : |t| \le T < \infty$.

Із цієї збіжності характеристичних функцій на підставі теореми Леві маємо твердження ЦГТ.

Більш загальне твердження центральної граничної теореми міститься нижче (без доведення).

Теорема 3. (Ляпунова). Нехай $\{\xi_n\}, n \ge 1$ – послідовність незалежних випадкових величин (не обов'язково однаково розподілених), для якої існує таке число $\delta > 0$, що $E|\xi_n|^{2+\delta} < +\infty \ \ \forall n \ge 1$. Позначимо $E\xi_n = a_n,\ \nu_n = \sum_{k=1}^{n} \xi_k,\ B_n{}^2 = D\nu_n \ \ \forall n \ge 1$. Нехай виконана умову Ляпунова: $L_n = \frac{1}{B_n{}^{2+\delta}} \sum_{k=1}^{n} E|\xi_k - a_k|^{2+\delta} \underset{n\to\infty}{\to} 0$. Тоді

$$\lim_{n\to\infty} P\left\{\frac{\nu_n - E\nu_n}{\sqrt{D\nu_n}} < x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad \forall x \in \mathbf{R} \tag{3}$$

Безпосереднім наслідком центральної граничної теореми є інтегральна теорема Мавра-Лапласа.

Теорема 4. (Інтегральна теорема Муавра-Лапласа). Нехай ν_n – кількість успіхів і серії n незалежних випробувань Бернуллі з ймовірністю успіху p та ймовірністю невдачі q, де $p+q=1$. Тоді

$$\lim_{n\to\infty} P\left\{\frac{\nu_n - np}{\sqrt{npq}} < x\right\} = \int_{-\infty}^{x} \frac{1}{2\pi} e^{-\frac{t^2}{2}} dt \text{ рівномірно по } x \in \mathbf{R}$$

Наслідок. (Довірчий інтервал (*confidence interval*) для ймовірності). Нехай p – невідома ймовірність здійснення деякої події A. Якщо провести серію із n незалежних випробувань, для якої обчислити кількість ν_n з'явлення події A, то з теореми Муавра-Лапласа випливає, що

$$P\left\{\left|\frac{\nu_n}{n} - p\right| < \varepsilon\right\} = P\left\{\left|\frac{\nu_n - np}{\sqrt{npq}}\right| < \varepsilon\sqrt{\frac{n}{pq}}\right\} \approx 2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right), \text{ де } \Phi_0(x) = \int_0^x \frac{1}{2\pi} e^{-\frac{t^2}{2}} dt.$$

Отже, якщо задатись деякою ймовірністю p_0, близькою до одиниці, то можна визначити по таблицях функції нормального розподілу таке значення ε, що $\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right) = \frac{p_0}{2}$. Тоді невідома ймовірність p події A з ймовірністю, наближено рівною p_0, міститься в інтервалі $\left[\frac{\nu_n}{n} - \varepsilon, \frac{\nu_n}{n} + \varepsilon\right]$: $P\left\{p \in \left[\frac{\nu_n}{n} - \varepsilon, \frac{\nu_n}{n} + \varepsilon\right]\right\} \approx p_0$.

2. Поняття про випадковий процес.

Найчастіше спостерігач має справу не просто із значеннями випадкової величини, а з деякою випадковою функцією, залежною від часу. Наприклад, величина напруги у електричній мережі коливається навколо деякого номінального значення і змінює випадковим чином своє значення у кожний момент часу. Те саме стосується багатьох інших природних чи фізичних процесів. Дамо точне означення випадкової функції.

Нехай $(\Omega, \Im, P)$ – деякий імовірнісний простір, а T – деяка підмножина множини дійсних чисел:і $T \subseteq \mathbf{R}$.

Означення 2. ***Випадковою функцією*** дійсного аргументу $t \in T$ називається функція $\xi(\omega, t): T \times \Omega \to \mathbf{R}$, яка при кожному значення аргументу t є випадковою величиною на імовірнісному просторі $(\Omega, \Im, P)$.

У позначенні випадкової функції, так само, як і в позначенні випадкової величини, залежність від елементарної події $\omega \in \Omega$ часто не записують: $\xi(\omega, t) = \xi_\omega(t) = \xi(t) = \xi_t$.

Означення 3. Якщо множина значень T відіграє роль часу, то випадкова функція $\xi(\omega, t)$ називається ***випадковим*** або ***стохастичним процесом***. Цю випадкову функцію називають ***випадковою послідовністю***, якщо множина T – дискретна. Якщо $T = \mathbf{R}^n$, то $\xi(\omega, t)$ називається ***випадковим полем***.

Отже, якщо зафіксувати значення елементарної випадкової події $\omega = \omega_0 \in \Omega$, то маємо ***невипадкову*** функцію $\xi(\omega_0, t)$, яка називається ***траєкторією*** або

реалізацією випадкового процесу. Якщо ж зафіксувати момент часу $t = t_0 \in T$, то матимемо випадкову величину $\xi(\omega, t_0)$, яку називають ***перерізом*** процесу. Таким чином, випадковий процес – це сукупність всіх його траєкторій або сукупність всіх його перерізів.

Приклад 1. Кількість $N(t)$ радіоактивних частинок, зареєстрованих лічильником за час t, – випадковий процес з цілими значеннями.

Приклад 2. Значення атмосферного тиску $P(t)$ в залежності від значення t – висоти над рівнем моря – є випадковим процесом з дійсними значеннями.

Приклад 3. Координати $X(t), Y(t), Z(t)$ деякої молекули рідини, що перебуває в процесі дифузії в просторі, в залежності від часу t – неперервні випадкові процеси з дійсними значеннями.

Означення 4. Функція $F(x,t) = P\{\omega \in \Omega : \xi(\omega, t) < x\} \overset{def}{=} P\{\xi(t) < x\}, x \in \mathbf{R}$ називається функцією ***розподілу випадкового процесу***.

Означення 5. ***Математичним сподіванням*** (***середнім***) випадкового процесу називається функція $m(t) = E\xi(t), \ t \in T$.

Означення 6. ***Дисперсією*** випадкового процесу називається функція $D(t) = D\xi(t) = E[\xi(t) - m(t)]^2, \ t \in T$.

Означення 7. ***Кореляційною функцією*** випадкового процесу називається функція $R(t_1, t_2) = E[(\xi(t_1) - m(t_1))(\xi(t_2) - m(t_2))], \ t_1, t_2 \in T$.

Приклад 4. Розглянемо для прикладу міркування, які привели до основного рівняння в теорії дифузії. (Ця задача розглядалась Максом та Фоккером).

Припустимо, що деяка частинка, що рухається вздовж прямої, стартує в нульовий момент часу із точки $x = 0$. Нехай у дискретні моменти часу $k\Delta t, \ k = 1, 2, \ldots$ під впливом незалежних випадкових поштовхів вона пересувається на величину h праворуч з ймовірністю p або ліворуч із ймовірністю $q = 1 - p$. Позначимо через $P(x,t)$ ймовірність того, що дана частинка опиниться у момент $t = n\Delta t$ (тобто через n кроків) в точці з абсцисою x. Очевидно, що функція $P(x,t)$ задовольняє різницевому рівнянню: $P(x, t + \Delta t) = pP(x - h, t) + qP(x + h, t)$ (4) із початковими умовами $P(0,0) = 1, \ P(x,0) = 0$ при $x \neq 0$. Спробуємо зробити граничний перехід при $h \to 0, \ \Delta t \to 0$. Із фізичних міркувань зрозуміло, що на параметри задачі мають бути накладені деякі обмеження, зокрема:

$$t = n\Delta t, \ x = nh, \ \frac{h^2}{\Delta t} \to 2D, \ \frac{p - q}{h} \to \frac{c}{D}, \tag{5}$$

де c та D – деякі сталі.

Віднімемо від обох частин рівності (4) $P(x,t)$ (врахувавши, що $P(x,t) = (p + q)P(x,t)$):

$$P(x, t + \Delta t) - P(x,t) = p[P(x - h, t) - P(x,t)] + q[P(x + h, t) - P(x,t)] \tag{6}$$

Припустимо, що функція $P(x,t)$ диференційована по t та двічі диференційована по x. Тоді

$$-P(x,t) + P(x - h, t) = -h\frac{\partial P(x,t)}{\partial x} + \frac{1}{2}h^2\frac{\partial^2 P(x,t)}{\partial x^2} + o(h^2),$$

$$P(x + h, t) - P(x,t) = h\frac{\partial P(x,t)}{\partial x} + \frac{1}{2}h^2\frac{\partial^2 P(x,t)}{\partial x^2} + o(h^2),$$

$$P(x,t+\Delta t)-P(x,t)=\Delta t\frac{\partial P(x,t)}{\partial t}+o(\Delta t)$$

Підставивши ці рівності у (6), матимемо:

$\Delta t\frac{\partial P(x,t)}{\partial t}+o(\Delta t)=-h(p-q)\frac{\partial P(x,t)}{\partial x}+\frac{1}{2}h^2\frac{\partial^2 P(x,t)}{\partial x^2}+o\left(h^2\right)$, звідки з допомогою (5) перейдемо до границі при $h\to 0,\ \Delta t\to 0$: $\frac{\partial P(x,t)}{\partial t}=-2c\frac{\partial P(x,t)}{\partial x}+D\frac{\partial^2 P(x,t)}{\partial x^2}$ – відоме рівняння дифузії, що носить назву рівняння Фоккера - Планка.

Один із класів випадкових процесів був розглянутий Н.Вінером – в теорію ймовірності він ввійшов під назвою ***вінерівського процесу*** або ***броунівського руху*** (хаотичний рух мікроскопічних частинок твердої речовини, що знаходяться в рідині або газі, під впливом теплового руху частинок середовища). На випадковий процес $w(t)$ були накладені природні наступні вимоги:

1. $w(0)=0;\ Ew(t)=0\ \forall t\in\mathbf{R}^+$
2. Функція розподілу різниці $w(t_0+t)-w(t_0)$ не залежить від початкового моменту t_0 : $P\{w(t_0+t)-w(t_0)<x\}=F_t(x)$ (однорідність процесу у часі).
3. $w(t)$ – процес із незалежними приростами, тобто прирости процесу за моменти часу , що не перетинаються, є незалежними, тобто $\forall n\ \forall t_1,\ldots,t_n\ w(t_1),\ w(t_2)-w(t_1),\ \ldots,\ w(t_n)-w(t_{n-1})$ – випадкові величини, незалежні у сукупності.
4. $E|w(t+h)-w(t)|^3=o(h), h\to 0$.

Зауваження 1. Властивості 2 та 3 визначають ОПНП – однорідний процес з незалежними приростами

Теорема 1. За припущень 1. – 4. вінерівський процес $w(t)$ має нормальний розподіл із параметрами 0 та $\sigma^2 t$, де σ^2 – деяка стала.

Доведення. Розглянемо $E[w(t+h)-w(t)]^2=\sigma^2(h), h>0$ – деяка функція, що залежить лише від h (із-за припущення 2). Покажемо, що насправді $\sigma^2(h)=\sigma^2\cdot h$ – лінійна по h функція. Нехай $h_1>0, h_2>0$. Тоді

$$\sigma^2(h_1+h_2)=E[w(t+h_1+h_2)-w(t)]^2=E[(w(t+h_1+h_2)-w(t+h_1))+(w(t+h_1)-w(t))]^2=$$
$$=\sigma^2(h_1)+\sigma^2(h_2)+2E[(w(t+h_1+h_2)-w(t+h_1))\cdot(w(t+h_1)-w(t))]=$$
$$=\sigma^2(h_1)+\sigma^2(h_2), \qquad (7)$$

в силу припущення 3. Крім того, $\sigma^2(h)$ – монотонна функція: $\sigma^2(h_1+h_2)\ge\sigma^2(h_1)$ і $\sigma^2(0)=0$. Відомо, що в класі монотонних функцій властивість адитивності (7) має лише лінійна функція, отже, $\sigma^2(h)=\sigma^2 h$. Позначимо $\varphi(z,t)=Me^{izw(t)}$ – характеристичну функцію вінерівського процесу. Тоді

$$\varphi(z,t+h)=Ee^{izw(t+h)}=Ee^{iz[w(t+h)-w(t)]+izw(t)}=\varphi(z,t)\cdot Ee^{iz[w(t+h)-w(t)]}.$$

Скористаємось розкладом в ряд Тейлора функції $e^{iz[w(t+h)-w(t)]}$ при $h\to 0$:

$$\varphi(z,t+h)=\varphi(z,t)\cdot E\left[1+\frac{iz[w(t+h)-w(t)]}{1!}-\frac{z^2[w(t+h)-w(t)]^2}{2!}+R\right]=$$

$$= \varphi(z,t)\left[1 - \frac{z^2}{2!}E[w(t+h) - w(t)]^2 + E(R)\right].$$

Оскільки $R = e^{iz\theta[w(t+h)-w(t)]} \cdot \frac{[w(t+h)-w(t)]^3}{3!}$, то в силу припущення 4, $E(R) = o(h)$.

Отже, при $h \to 0$:

$\varphi(z,t+h) = \varphi(z,t)\left[1 - \frac{z^2\sigma^2 h}{2} + o(h)\right] =$, тобто

$$\frac{\varphi(z,t+h) - \varphi(z,t)}{h} = -\varphi(z,t)\frac{z^2\sigma^2}{2} + \frac{o(h)}{h}\varphi(z,t), \quad \text{звідки} \quad \frac{\partial\varphi(z,t)}{\partial t} = -\frac{z^2\sigma^2}{2}\varphi(z,t).$$

Оскільки $\varphi(z,0) = 1$, то $\varphi(z,t) = e^{-\frac{z^2\sigma^2}{2}t}$ – характеристична функція розподілу $N(0,\sigma^2 t)$.

Зауваження 2. Можна показати, що траєкторії вінерівського процесу неперервні із ймовірністю 1: $P\{\omega : w(t,\omega) \in C([0,T])\} = 1$.

Зауваження 3. Можна показати, що траєкторії вінерівського процесу ніде не диференційовані із ймовірністю 1.

Теорема 2. Стандартний вінерівський процес має кореляційну функцію, рівну $R(t,s) = \min(t,s) \ \ \forall t,s \in [0,T]$.

Доведення. Для стандартного вінерівського процесу $\forall t \in \mathbf{R}: \ Ew(t) = 0, \ Ew^2(t) = t \ (\sigma^2 = 1)$. Припустимо, що $t > s$. Тоді

$R(t,s) = E[(w(t) - m(t))(w(s) - m(s))] = Ew(t)w(s) = E\left[(w(t) - w(s))w(s) + w^2(s)\right] =$

$= E[(w(t) - w(s))w(s)] + Ew^2(s) = Ew^2(s) = s = \min(t,s)$ (Тут ми використали незалежність величин $w(t) - w(s)$ та $w(s)$)

Одним із класів випадкових процесів, що описує так звані «рідкісні» події, є ***процес Пуассона***. Опишемо його. Припустимо, що деяка подія може відбуватись у випадкові моменти часу. Нехай $\xi(t)$ – кількість з'явлень цієї події на інтервалі $[0,t]$. Відносно процесу $\xi(t)$ зробимо наступні припущення.

1. Ймовірність k з'явлень цієї події на інтервалі $[t,t+\Delta t]$ залежить лише від k та Δt – ***стаціонарність*** процесу, зокрема ймовірність з'явлення цієї події принаймні раз на інтервалі $[t,t+\Delta t]$ покладемо рівною $\lambda\Delta t + o(\Delta t)$;
2. ймовірність з'явлення цієї події більше ніж один раз на інтервалі $[t,t+\Delta t]$ рівна $o(\Delta t)$ – ***ординарність*** процесу;
3. ймовірність k з'явлень цієї події на інтервалі $[t,t+\Delta t]$ не залежить від того, як і коли відбувались події раніше, – процес ***без післядії.*** Зокрема це означає взаємну незалежність з'явлення тієї чи іншої кількості подій протягом проміжків часу, що не перетинаються.

Теорема 3. З припущень 1. – 3. випливає, що процес $\xi(t)$ має розподіл Пуассона з параметром λ. Він називається ***процесом Пуассона.***

Доведення. Позначимо $P_k(t) = P\{\xi(t) = k\}, \ k = 0,1,\ldots$. Тоді

$P_0(t+\Delta t) = P_0(t)[1 - \lambda\Delta t + o(\Delta t)]$, тому $\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t) + \frac{o(\Delta t)}{\Delta t}P_0(t)$, звідки

в границі при $\Delta t \to 0$ одержимо $\frac{dP_0(t)}{dt} = -\lambda P_0(t)$, тому $P_0(t) = e^{-\lambda t}$ (оскільки $P_0(0) = 1$). Нехай тепер $k \geq 1$. Тоді за формулою повної ймовірності при $\Delta t \to 0$,

$P_k(t+\Delta t) = P_k(t)[1 - \lambda\Delta t + o(\Delta t)] + P_{k-1}(t)[\lambda\Delta t + o(\Delta t)] + o(\Delta t)$, звідки

$\frac{P_k(t+\Delta t) - P_k(t)}{\Delta t} = -\lambda P_k(t) + \lambda P_{k-1}(t) + \frac{o(\Delta t)}{\Delta t}[P_k(t) + P_{k-1}(t) + 1]$, тому для $k \geq 1$:

$$\frac{dP_k(t)}{dt} = -\lambda P_k(t) + \lambda P_{k-1}(t) \qquad (8)$$

із початковими умовами $P_0(0) = 1$ та $P_k(0) = 0$. Систему диференціальних рівнянь (8) розв'яжемо з допомогою заміни $P_k(t) = e^{-\lambda t} v_k(t), k \geq 1$. Початковими умовами для нових шуканих функцій $v_k(t)$, очевидно будуть наступні: $v_0(0) = 1, v_k(0) = 0, k \geq 1$. З (8) випливає, що

$e^{-\lambda t}\frac{dv_k(t)}{dt} - \lambda e^{-\lambda t} v_k(t) = -\lambda e^{-\lambda t} v_k(t) + \lambda e^{-\lambda t} v_{k-1}(t)$, або $\frac{dv_k(t)}{dt} = \lambda v_{k-1}(t), k \geq 1$. Тоді

$\frac{dv_1(t)}{dt} = \lambda v_0(t) = \lambda P_0(t) e^{\lambda t} = \lambda$, тобто $v_1(t) = \lambda t$. Аналогічно, $v_2(t) = \frac{(\lambda t)^2}{2}$ і взагалі,

$v_k(t) = \frac{(\lambda t)^k}{k!}, k \geq 1$. Остаточно маємо $P_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$, що й треба було довести.

Зауваження 4. Система припущень 1. – 3. виконується з великою точністю у багатьох фізичних явищах та технічних процесах, зокрема, їй відповідає кількість атомів при радіоактивному розпаді за певний проміжок часу або кількість космічних частинок, що потрапили на одиничну площадку за проміжок часу t , а також кількість деталей деякої системи, що вийшли з ладу протягом певного часу.

Література

1. Гнєденко Б. Курс теорії ймовірностей. – К: ВПЦ «Київський університет», 2010.–464 с.
2. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. – К.: Вища школа, 1979. – 408 с.
3. Карташов М.В. Імовірність, процеси, статистика : Посібник. – К.: Видавничо-поліграфічний центр 'Київський університет', 2007.– 494 с.
4. Іваненко Д.О., Сугакова О.В. «Самостійна робота студентів з курсу «Теорія ймовірностей і математична статистика»»
5. Шефтель З. Г. Теорія ймовірностей: Підручник. – К.: Вища школа, 1994. – 192 с.
6. Теорія ймовірностей: Збірник задач/ за ред. А.В. Скорохода. – К.: Вища школа, 1976. – 384 с.

Зміст

Лекція 1.	Предмет теорії ймовірностей. Випадкові події, їх ймовірності. Аксіоматика теорії ймовірностей. Елементи комбінаторики.	3
Лекція 2.	Властивості ймовірності. Умовні ймовірності, незалежні події. Формула повної ймовірності та формули Байеса.	9
Лекція 3.	Послідовності незалежних в*и*пробувань Бернуллі. Граничні теореми.	15
Лекція 4.	Випадкові величини та їх розподіли.	20
Лекція 5.	Випадкові вектори (Багатовимірні розподіли).	27
Лекція 6.	Числові характеристики випадкових величин.	32
Лекція 7.	Нерівність Чебишова. Закон великих чисел та його наслідки. Характеристична функція та її властивості.	38
Лекція 8.	Центральна гранична теорема. Поняття про випадковий процес.	45
	Література.	52