МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ І. І. МЕЧНИКОВА
ІНСТИТУТ МАТЕМАТИКИ, ЕКОНОМІКИ ТА МЕХАНІКИ

С. А. Щоголев

ОСНОВИ ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

НАВЧАЛЬНО-МЕТОДИЧНИЙ ПОСІБНИК

ОДЕСА
ОНУ
2015

УДК 519.2
ББК 22.17
Щ92

Рекомендовано до друку науково-методичною радою
ОНУ імені І. І. Мечникова.
Протокол № 5 від 19.06.2014 р.

Рецензенти:

В. Г. Попов – доктор фізико-математичних наук, професор, завідувач кафедри вищої математики Одеської національної морської академії;

А. В. Плотніков – доктор фізико-математичних наук, професор завідувач кафедри прикладної та обчислювальної математики і САПР Одеської державної академії будівництва та архітектури;

Ю. О. Григор'єв – кандидат фізико-математичних наук, доцент кафедри вищої і прикладної математики Одеського національного морського університету.

Щоголев С. А.

Щ92 **Основи теорії ймовірностей та математичної статистики:** навчально-методичний посібник / С. А. Щоголев. – Одеса : «Одеський національний університет імені І. І. Мечникова», 2015. – 206 с.

ISBN 978-617-689-131-4

Навчально-методичний посібник написано відповідно до навчальної програми дисципліни «Теорія ймовірностей та математична статистика» для студентів природничих спеціальностей.

Посібник містить основні поняття, методи, теореми та формули, багато розв'язаних типових задач, необхідні таблиці, а також завдання для самостійної роботи студентів.

УДК 519.2
ББК 22.17

ISBN 978-617-689-131-4

РОЗДІЛ І. ВИПАДКОВІ ПОДІЇ

1.1. Предмет теорії ймовірностей. Коротка історична довідка

Всі події, що відбуваються навколо нас, умовно можна поділити на дві групи. До першої групи відносяться так звані *детерміновані* події. Це такі події, про які завчасно можна стверджувати – відбудуться вони, чи не відбудуться. Наприклад, використовуючи закони класичної та небесної механіки, можна завбачити положення небесних тіл у обраний момент часу. До іншої групи відносяться *випадкові* події. Це такі події, які за наявності певного комплекса умов можуть відбутися, а можуть і не відбутися.

Приклади. 1. Нехай ми кидаємо монету. Чи можна визначити як саме вона впаде – «гербом» чи «решкою»? Напевно, якщо точно розрахувати рух монети, враховуючи всі фактори, що на його впливають (силу, з якою вона кинута, кут, під яким кинута, тощо), то можна передбачити випад, наприклад, «герба». Але зрозуміло, що за звичайних умов такий розрахунок практично неможливий. І практично ми не можемо завбачити ані випадіння «герба», ані «решки». Може відбутися й те, й інше. Таким чином, ми стверджуємо, що випадіння «герба» (також саме, як і «решки») – випадкова подія.

2. Нехай ми працюємо на комп’ютері – розв’язуємо якусь певну задачу, граємо в ігри, або користуємось мережею Інтернет. Чи можемо ми стверджувати, що комп’ютер буде працювати весь час, що нам потрібен? Очевидно, ні. Можлива наявність цілої низки факторів, які можуть зашкодити безперебійній роботі комп’ютера – може відмовити який-небудь вузол комп’ютера, може з’явитися вірус, який буде заважати, може відбутися роз’єднання з мережею, може вимикнути електрика тощо. Тому безперебійна робота комп’ютера (і взагалі будь-якого приладу) на протязі певного проміжку часу є подія випадкова.

Таким чином, кожна випадкова подія є наслідком багатьох факторів або невідомих нам причин, які впливають на подію. І врахувати цей вплив неможливо.

Разом з цим можна помітити наступне. Нехай, наприклад, ми кидаємо монету не один, а декілька разів поспіль. Тоді «герб» буде випадати приблизно в половині всіх кидань. Скажемо, якщо ми кидаємо монету 20 разів, то «герб» випадає, як правило, 8 – 12 разів. А ось якщо «герб» випаде всі 20 разів, що теоретично, взагалі кажучи, можливо, то одразу ж виникає думка, що це щось невипадкове. Наприклад, у відомій індійській кінострічці «Помста та закон» один з головних персонажів мав монету, на якій на обох сторонах був «герб». Зрозуміло, що «герб» випадав при кожному киданні монети, чим її власник успішно користувався. А за звичайних умов така ситуація практично неможлива.

Це свідчить про те, що, незважаючи на те, що появу кожної випадкової події окремо ми не можемо точно завбачити, при багаторазовому розгляді таких подій виникають певні закономірності. Вони називаються *ймовірнісними*. Дослідженням таких закономірностей займається теорія ймовірностей. Таким чином, предметом теорії ймовірностей є вивчення ймовірнісних закономірностей масових однорідних випадкових подій.

Результати та методи теорії ймовірностей широко використовуються в багатьох галузях природознавства і техніки: теорії надійності, теорії масового обслуговування, в теоретичній фізиці, астрономії, теорії стрільби, теорії помилок спостерігань, тощо. Теорія ймовірностей є також теоретичною основою для математичної та прикладної статистики, яка у свою чергу використовується при плануванні та організації виробництва, аналізі технологічних процесів, а також в гуманітарних науках – психології, соціології та ін.

Дуже важливу роль методи теорії ймовірностей та математичної статистики відіграють у геолого-географічних науках (це питання більш детально висвітлено у п. 3.1).

Перші роботи, у яких виникли основні поняття теорії ймовірностей, з'явилися у XV – XVI століттях як спроба побудови теорії азартних ігор і належать таким видатним вченим, як Б. Спіноза, Дж. Кардано, Г. Галілей. Серйозні вимоги з боку природознавства (теорія помилок спостерігань, теорія стрільби, проблеми статистики) привели до необхідності подальшого розвитку теорії ймовірностей та використання більш розвинутого математичного апарату (кінець XVII – початок XVIII ст.). Особливу значну роль тут відіграли роботи Б. Паскаля, П. Ферма, Х. Гюйгенса, К. Гауса, Я. Бернуллі, С. Пуассона, А. Муавра, П. Лапласа, Т. Байєса та ін.

Я. Бернуллі зробив перші теоретичні обґрунтування накопичених раніше фактів. Теорема, що ним доведена, отримала у подальшому назву «Закону великих чисел». З формально-аналітичного боку до цього ж напряму відноситься й робота М. І. Лобачевського, яку присвячено теорії помилок при вимірах на сфері.

Наприкінці XIX століття П. Л. Чебишов та його учні А. А. Марков та О. М. Ляпунов перетворили теорію ймовірностей у стрійну та послідовну математичну науку.

Подальшим розвитком теорія ймовірностей та випадкових процесів зобов'язана таким математикам, як С. М. Бернштейн, А. М. Колмогоров, О. Я. Хінчін, Б. В. Гнеденко, А. В. Скороход, В. С. Королюк, І. І. Гіхман, М. Й. Ядренко та ін.

1.2. Випадкові події та дії над ними

Означення. *Випадковою подією* називається така подія, яка при здійсненні певної сукупності умов може відбутися, а може й не відбутися.

Прикладами випадкових подій, як відмічалося у п. 1.1, може бути поява герба при киданні монети, поява карти певної масті при випадковому витяганні карти з колоди, влучання пулі в ціль при пострілі тощо.

Сукупність умов буде здійснена, якщо проведено певне випробування. Тому замість того, щоб казати «при здійсненні сукупності умов», будемо казати «внаслідок випробування». Тобто подія розглядатиметься як наслідок випробування.

Домовимось випадкові події позначати великими літерами латинського алфавіту: A, B, C, …

Означення. Подія називається *достовірною*, якщо внаслідок випробування вона обов'язково трапиться.

Приклад. Нехай з скриньки, в якій містяться лише білі кулі, навмання витягається одна куля. Подія, яка полягає в тому, що ця куля біла, є достовірною.

Означення. Подія називається *неможливою*, якщо внаслідок випробування вона не може трапитися.

Наприклад, поява з скриньки, де тільки білі кулі, кулі будь якого іншого кольору є неможлива подія.

Домовимось достовірну подію позначати літерою E, а неможливу – літерою O. Фактично достовірна та неможлива подія не є випадковими. Але, як ми побачимо далі, їх можна розглядати як певні граничні випадки випадкових подій.

Означення. Випадкова подія $\overline{A}$ називається *протилежною* випадковій події A, якщо вона полягає в тому, що внаслідок випробування подія A на трапиться.

Наприклад, нехай з колекції зразків різних мінералів навмання обирається один зразок. Якщо подія A полягає в тому, що цей зразок польовий шпат, то подія $\overline{A}$ – в тому, що цей зразок не є польовим шпатом.

Означення. *Добутком* AB випадкових подій A і B називається подія, яка полягає в тому, що внаслідок випробування трапилася як подія A, так і подія B.

Тобто подія AB полягає у сумісній появі подій A і B.

Приклад. Нехай з колоди карт навмання витягається одна карта. Якщо подія A – ця карта пікової масті, а подія B – ця карта дама, то подія AB – ця карта пікова дама (водночас і піка, і дама).

Означення. Події A і B називаються *несумісними*, якщо AB – подія неможлива; $AB = O$.

Тобто несумісні події – це такі події, які в одному і тому ж випробування виключають одна іншу. Разом вони відбутися не можуть.

Приклад. У коробці лежать шашки, чорні та білі. Навмання витягається одна шашка. Якщо подія A – ця шашка біла, а подія B – ця шашка чорна, то A і B – несумісні події.

Якщо події A і B не є несумісними, то вони називаються *сумісними*.

Означення. *Сумою* $A+B$ випадкових подій A і B називається подія, яка полягає в тому, що внаслідок випробування відбудеться хоча б одна з подій A і B.

Тобто трапиться або тільки подія A, або тільки подія B, або події A і B разом.

Приклад. Нехай здійснюється два постріли в мішень. Якщо подія A – влучання при першому пострілі, подія B – влучання при другому пострілі, то $A+B$ – влучання або тільки при першому пострілі, або тільки при другому пострілі, або в обох пострілах. Тобто відбулося хоча б одне влучання.

Сформулюємо деякі властивості дій над подіями.

1. Подія, протилежна достовірній, є неможливою, а подія, протилежна неможливій, є достовірною, тобто $\overline{E} = O$, $\overline{O} = E$.

2. Подія, протилежна події, яка протилежна події A, є подія A, тобто $\overline{\overline{A}} = A$.

3. Добуток події A на її протилежну є неможлива подія, тобто $A\overline{A} = O$. Одна й та ж подія не може водночас і трапитись і не трапитись (тут виключаються поетичні метафори типу «речка движется и не движется», «песня слышится и не слышится»).

4. Сума події A та її протилежної є подія достовірна, тобто $A + \overline{A} = E$. Інакше кажучи, будь-яка подія завжди має або відбутися, або не відбутися. Персонаж відомого мультфільму Вінні-Пух помітив, що будь-яка річ (крім меда) завжди або є, або її нема. Тому жартівливо можна назвати цю властивість теоремою Вінні-Пуха.

5. Множення подій є операція комутативна, тобто $AB = BA$. Це безпосередньо випливає з означення добутку подій: AB і BA – одна й та ж сумісна поява подій A і B.

6. Додавання подій є також операція комутативна, тобто $A+B = B+A$. Це також безпосередньо випливає з означення суми подій.

7. Множення та додавання подій є операції асоціативні, тобто: $A(BC) = (AB)C$, $A+(B+C) = (A+B)+C$. Це також безпосередньо випливає з означення добутку та суми подій.

8. Має місце дистрибутивний закон, тобто $A(B+C)=AB+AC$.

А ця властивість вже вимагає доведення. Дійсно, подія $B+C$ полягає в здійсненні хоча б однієї з подій B або C. Тому подія $A(B+C)$ полягає в здійсненні хоча б однієї з подій B або C, а також обов'язковому здійсненні події A. А це означає, що подія A має відбутися або разом тільки з подією B, або разом тільки з подією C, або як з подією B, так і з подією C водночас. Тобто має відбутися хоча б одна з подій AB або AC. А це й є сума $AB+AC$.

І навпаки, сума $AB+AC$ полягає в тому, що подія A здійснилася або разом з подією B, або разом з подією C, або з ними обома водночас. А це означає, що подія A здійснилася разом хоча б з однією з подій B і C. А це й є $A(B+C)$.

Легко бачити, що властивості 5 – 8 дій над подіями аналогічні діям додавання та множення дійсних чисел a, b, c: $ab=ba$, $a+b=b+a$, $a(b+c)= =ab+ac$.

9. Добуток події A на достовірну подію є подія A: $AE=EA=A$.

Дійсно, добуток AE або EA полягає у сумісному здійсненні подій A і E. Але подія E достовірна, вона трапиться обов'язково, тому подія AE або EA трапиться тоді і тільки тоді, коли трапиться подія A, а це й означає, що $AE=EA=A$.

10. Добуток події A на неможливу подію є подія неможлива: $AO=OA=O$.

Дійсно, добуток AO або OA полягає у сумісному здійсненні події A і неможливої події, але неможлива подія не може відбутися, отже, не може відбутися й AO або OA, а це й означає, що $AO=OA=O$.

11. Сума події A і неможливої події є подія A, тобто $A+O=O+A=A$.

Дійсно, сума $A+O$ або $O+A$ полягає в здійсненні або події A або неможливої. Але оскільки неможлива подія не може трапитися, то подія $A+O$ або $O+A$ трапиться тоді і тільки тоді, коли трапиться подія A, а це й означає, що $A+O=O+A=A$.

Властивості 9 – 11 показують, що в діях над подіями неможлива подія O та достовірна подія E є певними аналогами нуля та одиниці в діях над дійсними числами. А саме $a\cdot 0=0$, $a\cdot 1=a$, $a+0=a$. Але в той же час є й відмінність. В діях над дійсними числами $a+1\neq 1$, якщо тільки $a\neq 0$. Тем не менш в діях над подіями маємо наступну властивість.

12. Сума події A і достовірної події є подія достовірна: $A+E=E+A=E$.

Дійсно, сума $A+E$ або $E+A$ полягає в здійсненні хоча б однієї з подій – події A або достовірної. Але достовірна подія трапиться обов'язково, тому обов'язково трапиться і $A+E$ (як і $E+A$). А це й означає, що $A+E= =E+A=E$.

13. Подія, протилежна добутку подій A і B, є сума подій, протилежних подіям A і B: $\overline{AB}=\overline{A}+\overline{B}$.

Дійсно, розглянемо спочатку подію $\overline{AB}$. Вона полягає в тому, що події A і B не трапилися разом. А тоді або не трапилася подія A, або подія B, або вони обидві не трапилися. В свою чергу це означає, що трапилася або подія $\overline{A}$, або подія $\overline{B}$, або разом події $\overline{A}$ і $\overline{B}$. А це й є сума $\overline{A}+\overline{B}$.

Тепер навпаки, розглянемо подію $\overline{A}+\overline{B}$. Вона полягає в тому, що трапилися або подія $\overline{A}$, або подія $\overline{B}$, або події $\overline{A}$ і $\overline{B}$ разом. А це означає, що не трапилася або подія A, або подія B, або вони обидві не трапилися. В усякому разі події A і B не трапилися разом, а це й означає подію, протилежну до AB, тобто $\overline{AB}$. Таким чином, подія $\overline{AB}$ трапляється тоді і тільки тоді, коли трапляється подія $\overline{A}+\overline{B}$. Це й означає, що $\overline{AB}=\overline{A}+\overline{B}$.

14. Подія, протилежна сумі подій A і B, є добуток подій, протилежних подіям A і B: $\overline{A+B}=\overline{A}\overline{B}$.

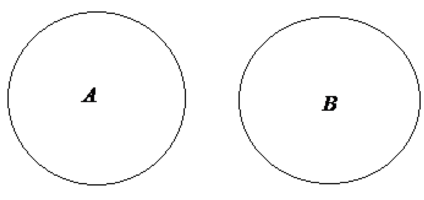
Дійсно, розглянемо спочатку подію $\overline{A+B}$. Вона полягає в тому, що подія, що полягає в здійснені хоча б однієї з подій A і B, не трапилася. А це означає, що не трапилася жодна з подій A і B. Тобто разом трапилися події $\overline{A}$ і $\overline{B}$, а це й є подія $\overline{A}\overline{B}$.

Тепер розглянемо подію $\overline{A}\overline{B}$. Вона полягає в тому, що трапилися разом події $\overline{A}$ і $\overline{B}$, тобто не трапилися як подія A, так і подія B. А це означає, що не трапилася навіть хоча б одна з подій A і B, тобто трапилася подія $\overline{A+B}$.

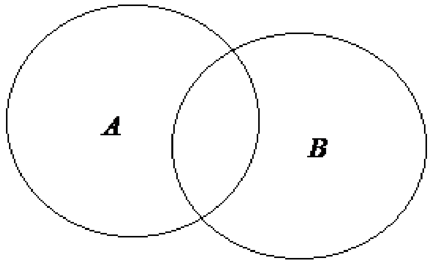
Таким чином подія $\overline{A+B}$ трапляється тоді і тільки тоді, коли трапляється $\overline{A}\overline{B}$. Це й означає, що $\overline{A+B}=\overline{A}\overline{B}$.

Властивості 13, 14 називаються *законами де Моргана.*

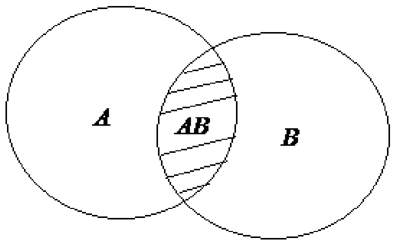
Діям над подіями можна надати досить наглядні інтерпретації. Розглянемо геометричну інтерпретацію. Домовимось випадкові події A і B зображувати кругами на площині:



Події A і B будуть сумісними, якщо ці круги матимуть непорожній перетин:

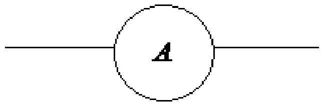


Добутку AB подій A і B відповідатиме саме цей перетин:

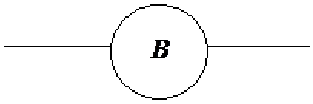


Сумі подій A і B відповідатиме об'єднання кругів.

Інша інтерпретація суми та добутку подій пов'язана з електричними схемами. Припустимо, що є електричні лампи A і B. І нехай подія A полягає в тому, що лампа A пропускає електричний струм, а подія B – що лампа B пропускає електричний струм.

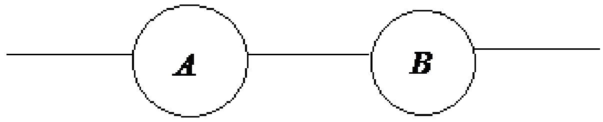


подія A



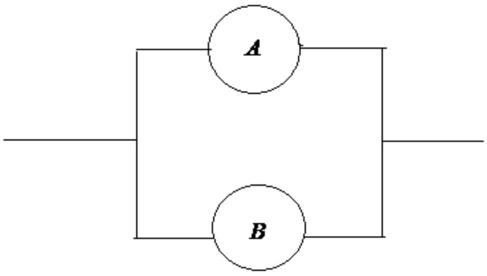
подія B

Якщо лампи A і B включено до електричного ланцюга послідовно, то ланцюг пропускатиме струм тоді і тільки тоді, коли пропускають струм і лампа A, і лампа B. Тобто коли трапляється добуток подій A і B:



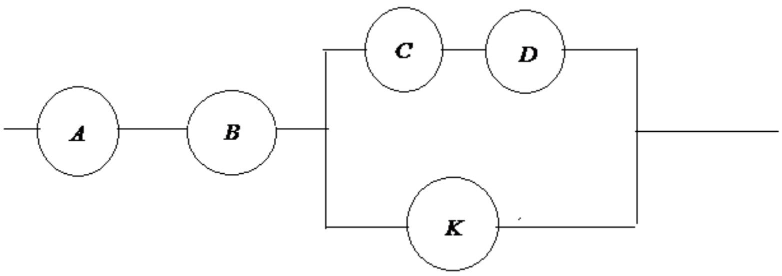
подія AB

Якщо лампи A і B включено паралельно, то ланцюг пропускатиме струм тоді і тільки тоді, коли пропускає струм хоча б одна з ламп A і B. Тобто коли трапляється сума подій A і B:



подія $A+B$

Приклад. Задано електричного ланцюга:



Записати подію, яка полягає в тому, що цей ланцюг пропускає струм.

Очевидно, що ця подія є: $AB(CD+K)$.

1.3. Класичне означення ймовірності події

Почнемо з розгляду прикладу. Нехай є скринька, у якій містяться 10 куль, однакових за розміром та матеріалом, але різних кольорів: 6 чорних, 3 білих і одна червона. І навмання зі скриньки вибирається одна куля. Ця куля може з'явитися чорною, може білою, а може й червоною. Тобто всі ці варіанти можливі. Але очевидно, що можливості у них різні. Дійсно, можливість витягнути чорну кулю більша, ніж білу, і ще більша, ніж червону. Таким чином, ми бачимо, що можливі події можуть мати різні можливості для здійснення. Чи можна надати числову характеристику можливості події? Так, можна. Цю характеристику називають ймовірністю події. Розглянемо це питання детальніше. Нехай подія A – поява чорної кулі, подія B – поява білої кулі, подія C – поява червоної кулі. Кожен з можливих наслідків випробування (випробування полягає у взятті кулі зі скриньки) назвемо *елементарним наслідком*. Елементарні наслідки будемо позначати $\omega_1, \omega_2, ..., \omega_{10}$. Всі ці наслідки рівноможливі, і один з них після випробування обов'язково з'явиться. Елементарні наслідки, в яких подія, що нас цікавить, відбувається, назвемо наслідками, які *сприяють* появі цій події (або *сприятливі наслідки*). Будемо вважати, що появі події A сприяють наслідки $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$, події B – наслідки ω_7, ω_8, ω_9, а події C – наслідок ω_{10}. Таким чином, подія A відбувається, якщо при випробуванні виникає один з елементарних наслідків $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$. У цьому сенсі подія A розділяється на декілька елементарних подій ($\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$). Відповідно подія B розділяється на елементарні події ω_7, ω_8, ω_9, а подія C складається тільки з однієї елементарної події ω_{10}. А ось елементарна подія на інші події не розділяється.

Очевидно, що чим більше рівноможливих елементарних наслідків сприяє появі події, тим більше можливостей для появи цієї події. У нашому прикладі найбільшу можливість має подія A (6 елементарних наслідків), а найменшу – подія C (1 елементарний наслідок). У якості числової характеристики цієї можливості (ймовірності) можна взяти відношення числа елементарних наслідків, які сприяють появі події, до загального числа всіх рівноможливих елементарних наслідків. Для події A це буде $6/10$, для події B – $3/10$, для події C – $1/10$.

Дамо тепер загальне означення ймовірності події.

Означення. *Ймовірністю* події A називається відношення числа рівноможливих елементарних наслідків випробування, які сприяють появі події A, до загального числа всіх рівноможливих елементарних наслідків випробування.

Ймовірність події A позначається $P(A)$ (P – перша буква англійського слова *probability* – ймовірність). Таким чином:

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

де m – число елементарних наслідків, які сприяють появі події A, а n – загальне число всіх рівноможливих елементарних наслідків.

Наведене означення називається *класичним* означенням ймовірності події. З нього випливають наступні її властивості:

1. Ймовірність достовірної події дорівнює одиниці: $P(E)=1$.

Дійсно, якщо подія достовірна, то всі можливі елементарні наслідки сприяють її появі, тобто $m=n$, отже

$$P(E)=\frac{m}{n}=\frac{n}{n}=1.$$

2. Ймовірність неможливої події дорівнює нулю: $P(O)=0$.

Дійсно, якщо подія неможлива, то жоден з елементарних наслідків випробування не сприяє появі події, тобто $m=0$, отже

$$P(O)=\frac{m}{n}=\frac{0}{n}=0.$$

3. Ймовірність випадкової події є число, яке лежить між нулем та одиницею: $0<P(A)<1$.

Дійсно, якщо подія випадкова (не є достовірною та не є неможливою), то $0<m<n$, отже $0<m/n<1$, тобто $0<P(A)<1$.

Таким чином, ймовірність будь якої події задовольняє подвійну нерівність:

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

1.4. Деякі поняття комбінаторики

Для розуміння подальшого матеріалу нам будуть потрібні деякі поняття комбінаторики. *Комбінаторика* – це розділ математики, який вивчає закономірності, що виникають при розгляді комбінацій, які можна скласти з елементів будь-якої природи. Наведемо у якості прикладу типово комбінаторну задачу. У студентській групі навчаються 25 осіб. Треба з них виділити трьох делегатів на студентську конференцію. Скількома способами це можна зробити?

Основним принципом комбінаторики є наступний. Якщо елемент a можна вибрати n способами, а елемент b можна вибрати m способами, то пару елементів (a,b) можна вибрати nm способами. Наприклад, якщо з пункту A до пункту B ведуть n різних шляхів, а з пункту B до пункту C ведуть m різних шляхів, то з пункту A до пункту C ведуть nm різних шляхів.

Під час розв’язування комбінаторних задач доводиться розглядати скінченні множини, що складаються з означених елементів. Залежно від умови задачі розглядаються множини, в яких істотним є або порядок елементів, або їх склад, або перше і друге водночас. Такі скінченні множини (сполуки) дістали назву: переставлення, розміщення, сполучення.

Нехай є множина A, яка складається з n елементів довільної природи.

Сполуки, що складаються з всіх елементів множини A, і відрізняються лише порядком їх перелічення, називаються *переставленням* цих елементів. Кількість переставлень з n елементів позначається P_n.

<u>Приклад</u>. Нехай $A=\{1,2,3\}$. Розглянемо всі сполуки, що складаються з всіх елементів множини A:

$\{1,2,3\}$, $\{1,3,2\}$, $\{2,1,3\}$, $\{2,3,1\}$, $\{3,1,2\}$, $\{3,2,1\}$.

Вийшло 6 сполук, таким чином $P_3=6=3!$. Легко зрозуміти, що $P_1=1=1!$, $P_2=2=2!$.

Розглянемо множину, що складається з 4-х елементів: $\{1,2,3,4\}$. Знайдемо P_4. Скільки переставлень можна скласти з 4-х елементів? До них увійдуть 6 переставлень, які мають на 1-му місці 1 (відрізняються тільки трьома подальшими елементами), 6 переставлень, які мають на 1-му місці 2, 6 переставлень, які мають на 1-му місці 3, і 6 переставлень, які мають на 1-му місці 4. Всього вийшло $6\cdot 4=1\cdot 2\cdot 3\cdot 4=4!=24$ переставлень. Тобто $P_4=4!$. Методом математичної індукції нескладно встановити загальну формулу:

$$P_n=n!.$$

<u>Приклад</u>. Скількома способами можна розмістити 12 осіб за столом, біля якого поставлено 12 стільців?

Число способів дорівнює $P_{12}=12!=479\,001\,600$.

Нехай тепер з елементів множини A утворюються всі можливі сполуки, що містять k елементів ($0\le k\le n$). При цьому не враховується порядок перелічення елементів у кожній з сполук. Тобто сполуки, що відрізняються тільки порядком перелічення елементів, не розрізняються. Такі сполуки називають *сполученнями з* n *по* k. Число всіх таких сполук позначається C_n^k. Можна довести формулу:

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

Зокрема, $C_n^0=C_n^n=1$, $C_n^1=C_n^{n-1}=n$. Корисними є також формули:

$$C_n^k=C_n^{n-k},$$

$$C_n^0+C_n^1+C_n^2+\ldots+C_n^n=2^n.$$

<u>Приклад.</u> Студент прийшов до бібліотеки та запросив 7 книжок. А йому можуть видати лише 4 книжки, але будь-яких з цих семи. Скількома способами студент може вибрати ці 4 книжки?

Очевидно, що порядок перелічення цих книжок у наборі з 4 книжок не суттєвий, тому шукана кількість способів дорівнює числу сполучень з 7 по 4:

$$C_7^4 = \frac{7!}{4!3!} = \frac{5 \cdot 6 \cdot 7}{3!} = \frac{5 \cdot 6 \cdot 7}{6} = 5 \cdot 7 = 35.$$

Число сполучень C_n^k є число способів, якими можна вибрати k елементів з множини, що містить n елементів.

Будемо тепер з елементів множини A знову утворювати всі можливі сполуки, які містять k елементів, але цього разу з урахуванням порядку перелічення елементів у кожній з сполук. Тобто сполуки, які складаються з одних і тих же елементів і розрізняються тільки порядком їх перелічення, тепер вважаються різними. Такі сполуки називаються *розміщеннями з n по k*. Число всіх таких сполук позначається A_n^k. Легко зрозуміти, що це число може бути знайдено за формулою:

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

Приклад. Скільки існує 6-значних телефонних номерів, всі цифри яких різні?

Очевидно, що ці телефонні номери – сполуки, які утворено з елементів 10-елементної множини, і містять 6 елементів, причому порядок перелічення елементів у кожній сполуці суттєвий (номери, що складаються з одних і тих же цифр, але розрізняються порядком їх слідування, очевидно, різні). Отже число всіх таких номерів є число розміщень з 10 по 6:

$$A_{10}^6 = \frac{10!}{4!} = 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 = 151\ 200.$$

1.5. Задачі на безпосередній підрахунок ймовірностей

Задача 1. Дитина грає з 5 картками, на яких написано букви О, П, Р, С, Т (на кожній картці одна з цих букв). Яка ймовірність того, що, викладаючи ці картки у випадковому порядку, дитина отримає слово «СПОРТ»?

Скористаємось класичним означенням ймовірності. Загальна кількість всіх рівноможливих елементарних наслідків дорівнює числу переставлень з 5 елементів, тобто $P_5 = 5! = 120$. А елементарний наслідок, що сприяє події, тільки один. Таким чином, $m = 1$, $n = 120$, отже

$$P = \frac{m}{n} = \frac{1}{120}.$$

Задача 2. Набираючи номер телефону, абонент забув дві останні цифри і набрав їх навмання, пам'ятаючи тільки, що вони непарні та різні. Знайти ймовірність того, що номер набрано правильно.

Позначимо через A подію, яка полягає в тому, що набрано потрібні цифри. Ці дві цифри набираються з 5-елементної множини $\{1,3,5,7,9\}$, до того ж суттєвим є порядок їх слідування. Отже число всіх елементарних наслідків дорівнює

$$A_5^2 = \frac{5!}{3!} = 4 \cdot 5 = 20.$$

А сприяє події A тільки один наслідок. Таким чином, $m = 1$, $n = 20$, отже,

$$P(A) = \frac{1}{20}.$$

Задача 3. На одній з кліток шахівниці стоїть біла тура. На якусь іншу клітку навмання ставиться чорна тура. Знайти ймовірність того, що вона не потрапить під бій білої.

Кількість кліток, які контролює тура, не залежить від клітки, на якій вона стоїть – у будь-якому випадку вона контролює 14 кліток (7 по горизонталі та 7 по вертикалі). Вільних кліток для чорної тури залишається 63, а з них тих, що не контролюється білою турою – 63-14=49. Отже, $m = 49$, $n = 63$, таким чином:

$$P = \frac{49}{63} = \frac{7}{9}.$$

Задача 4. Галя, Марійка, Василь, Петро, Оксана та Наталя прийшли на День народження до Ганни. Всі вони сідають на 7 стільців біля круглого столу. Знайти ймовірність того, що Василь опиниться поряд з Ганною.

Незалежно від стільця, який займе Ганна, вільними залишаються 6 стільців для 6 гостей. Нехай A – подія, яка полягає в тому, що Василь опиниться поряд з Ганною. Загальна кількість всіх елементарних наслідків це число способів, якими можуть сісти 6 гостей Ганни, а це є число переставлень з 6 елементів, тобто $P_6 = 6!$. Число елементарних наслідків, які сприяють події A – $2 \cdot 5!$. Дійсно, Василь може опинитися ліворуч від Ганни, а може праворуч; у кожному з цих випадків решта гостей можуть сісти будь-яким чином, а число таких варіантів 5!. Таким чином,

$$P(A) = \frac{2 \cdot 5!}{6!} = \frac{2}{6} = \frac{1}{3}.$$

Цю задачу можна було б розв'язати і простіше: з 6 місць, що залишилося, поряд з Ганною є тільки 2. Отже шукана ймовірність дорівнює $2/6$.

Якщо кількість осіб, що сідають біля круглого столу, дорівнює n, то аналогічно доводиться, що ймовірність того, що дві конкретні особи опиняться поряд, дорівнює $2/(n-1)$.

<u>Задача 5</u>. Кидаються три монети. Знайти ймовірність того, що «герб» випаде точно на 2 монетах.

Розглянемо всі рівноможливі елементарні наслідки випробування. Випадіння «герба» позначатимемо як Г, а випадіння «решки» – як Р:

ГГГ, ГГР, ГРГ, ГРР, РГГ, РГР, РРГ, РРР.

Таким чином, $n=8$. А елементарних наслідків, що сприяють нашій події – 3: ГГР, ГРГ, РГГ. Тобто $m=3$, отже

$$P=\frac{3}{8}.$$

<u>Задача 6</u>. Навмання обрана кістка доміно з'явилася не дублем. Знайти ймовірність того, що другу, навмання обрану кістку, можна приставити до першої.

Всього кісток доміно 28, з них дублів 7. Якщо обрана кістка не дубль, наприклад 2-3, то до неї з 27 кісток, що залишилося, можна приставити 12, а саме: 0-2, 1-2, 2-2, 4-2, 5-2, 6-2, 0-3, 1-3, 3-3, 4-3, 5-3, 6-3. Таким чином, $n=27$, $m=12$, отже

$$P=\frac{12}{27}=\frac{4}{9}.$$

<u>Задача 7</u>. 9 мишей біжать до трьох відкритих кліток з кормом. Кожна миша з рівними можливостями може потрапити до будь-якої з цих трьох кліток. Знайти ймовірність того, що у кожній клітці опиниться по 3 миші.

Кожна миша має три рівноможливі варіанти потрапляння в клітки. Згідно з основним принципом комбінаторики (див. п. 1.4) всіх варіантів потрапляння всіх 9 мишей буде $3\cdot3\cdot3\cdot3\cdot3\cdot3\cdot3\cdot3\cdot3=3^9$. Це й буде число всіх елементарних наслідків випробування. Для знаходження числа сприятливих наслідків підрахуємо спочатку кількість способів, якими можна обрати трьох мишей з 9, що є (у першій клітці мають бути три миші). Це буде C_9^3. З 6 мишей, що залишилося, три мають бути у другій клітці, їх можна обрати C_6^3 способами. І нарешті останні три миші потрапляють в третю клітку вже єдиним способом. Таким чином, $n=3^9$, $m=C_9^3\cdot C_6^3\cdot 1$, отже,

$$P=\frac{m}{n}=\frac{C_9^3\cdot C_6^3}{3^9}=\frac{1680}{19683}\approx 0{,}085.$$

Задача 8 (про вибірку). У скриньці містяться 15 куль, з яких 9 білих та 6 чорних. Навмання з скриньки виймаються 10 куль. Знайти ймовірність того, що серед цих 10 рівно 7 білих та 3 чорних.

Загальне число всіх елементарних наслідків дорівнює числу способів, якими можна вибрати 10 куль з 15, тобто $n = C_{15}^{10}$. Обчислимо число сприятливих наслідків. До цього числа увійде кількість способів, якими можна вибрати 7 білих куль з 9, що їх є, тобто C_9^7. На кожен з цих способів прийдеться ще C_6^3 способів, якими можна вибрати 3 чорних кулі з 6, що їх є. Згідно з основним принципом комбінаторики маємо: $m = C_9^7 \cdot C_6^3$. Таким чином:

$$P = \frac{m}{n} = \frac{C_9^7 \cdot C_6^3}{C_{15}^{10}} = \frac{\dfrac{9!}{7!2!} \cdot \dfrac{6!}{3!3!}}{\dfrac{15!}{10!5!}} = \frac{9!\,6!\,10!\,5!}{7!\,2!\,3!\,3!\,15!} = \frac{20}{429} \approx 0{,}0466.$$

Сформулюємо цю задачу та її розв'язок у загальному випадку. Нехай є N предметів, з них n предметів мають властивість s. Навмання вибираються M предметів. Знайти ймовірність того, що серед них рівно m мають властивість s.

Шукана ймовірність дорівнює:

$$P = \frac{C_n^m \cdot C_{N-n}^{M-m}}{C_N^M}\ .$$

Задача 9. З колоди в 52 карти навмання вибираються 3 карти. Знайти ймовірність того, що це будуть трійка, сімка і туз.

Загальна кількість всіх елементарних наслідків дорівнює числу способів, якими можна вибрати 3 карти з 52, тобто $n = C_{52}^3$. Трійку можна вибрати C_4^1 способами, така ж кількість способів вибору сімки та туза. Згідно з основним принципом комбінаторики $m = C_4^1 \cdot C_4^1 \cdot C_4^1 = 4^3 = 64$. Таким чином:

$$P = \frac{m}{n} = \frac{64}{C_{52}^3} = \frac{64 \cdot 6}{50 \cdot 51 \cdot 52} \approx 0{,}0029.$$

Задача 10. Вісім різних книжок поставлено навмання на одній полиці. Знайти ймовірність того, що дві певні книжки опиняться поряд.

Загальна кількість елементарних наслідків дорівнює $P_8 = 8!$. Знайдемо число сприятливих наслідків. Дві книжки можна поставити поряд $2! = 2$ способами, тоді решту книжок можна розставити $6!$ способами. Крім того маємо 7 можливих положень цих двох книжок вздовж полиці. Отже $m = 7 \cdot 2! \cdot 6!$, і таким чином:

$$P = \frac{m}{n} = \frac{7 \cdot 2! \cdot 6!}{8!} = \frac{1}{4}.$$

Задача 11. Знайти помилку у розв’язанні наступної жартівливої задачі: яка ймовірність, вийшовши в певний час доби на Дерибасівську, зустріти там справжнього живого динозавра?

Розв’язання: всього елементарних наслідків випробування 2 – або зустріти динозавра, або не зустріти. З них сприятливий тільки один – коли зустріти. Отже ймовірність дорівнює $1/2$ (така ж сама, як ймовірність появи герба при киданні монети).

1.6. Статистичне означення ймовірності

Класичне означення ймовірності події має суттєві обмеження. По-перше, припускається, що кількість елементарних наслідків випробування скінченна. В той же час на практиці зустрічаються випробування, де кількість наслідків нескінченна, наприклад, у геометричних задачах. І навіть у тих випадках, коли кількість елементарних наслідків скінченна, вона може бути настільки великою, що практично підрахувати її неможливо, а також неможливо підрахувати число сприятливих наслідків (наприклад, у статистичній фізиці). По-друге, часто буває неможливо подати наслідок випробування у вигляді сукупності елементарних подій. Ще складніше вказати підстави, за яких можна вважати ці елементарні події рівноможливими. За цими причинами поряд з класичним означенням ймовірності, користуються іншими (аксіоматичне, геометричне, статистичне). Тут ми познайомимось зі статистичним означенням.

Нехай ми проводимо не одне випробування, а цілу серію випробувань. Число цих випробувань позначимо як N. Припустимо, що в частині цих випробувань з’явилася подія A, число цих випробувань позначимо як M .

Означення. *Відносною частотою* події A називається відношення числа випробувань, у яких з’явилася подія A, до загального числа випробувань.

Таким чином, відносна частота $\nu(A)$ події A визначається формулою:

$$\nu(A) = \frac{M}{N},$$

де M – число появ події A, N – загальне число випробувань.

Приклад. Монета кидається 40 разів, з них 24 рази випав герб. Відносна частота появи герба:

$$\nu = \frac{24}{40} = \frac{3}{5}.$$

Очевидно, що якщо змінюється загальне число випробувань N, то, взагалі кажучи, буде змінюватись і число M появ події A, тобто $M = M(N)$.

Відповідно і відносна частота, взагалі кажучи, залежить від N : $\nu(A)=\nu_N(A)$. Але спостерігання показали, що якщо в однакових умовах проводять експерименти, в кожному з яких число випробувань дуже велике, то відносна частота демонструє властивість стійкості, тобто в різних експериментах відносна частота змінюється незначно (тим менше, чим більше проведено випробувань). І ця відносна частота, як правило, залишається близькою до певної сталої величини. З'ясувалося, що це число – не що інше, як ймовірність події.

Таким чином, якщо встановлено відносну частоту події, то цю частоту можна прийняти за наближене значення ймовірності. Можна сказати, що відносна частота події є певним статистичним аналогом ймовірності цієї події.

Наведемо наступний приклад. Проводилися експерименти з киданням монети і фіксувалося число появ герба. Результати експериментів наведено в наступній таблиці:

Число кидань монети	Число появ герба	Відносна частота появ герба
4 040	2 048	0,5069
12 000	6 019	0,5016
24 000	12 012	0,5005

Ми бачимо, що при збільшенні числа кидань відносна частота появи герба стає все ближчою до числа 0,5, а це як раз ймовірність появи герба при одноразовому киданні монети. Цей факт є підставою для наступного (статистичного) означення ймовірності події.

Означення. *Ймовірністю* події A називається границя відносної частоти події A при необмеженому збільшенні числа випробувань. Тобто:

$$P(A)=\lim_{N\to\infty}\nu_N(A)=\lim_{N\to\infty}\frac{M(N)}{N}.$$

Очевидно, що таке означення ймовірності також має певні недоліки. По-перше, необхідна можливість, хоча б принципова, проведення необмеженого числа випробувань. По-друге, потрібні умови, що забезпечують стійкість відносних частот появи події A в різних серіях достатньо великої кількості випробувань. І навіть якщо такі умови є, ми не можемо стверджувати, що при збільшенні кількості випробувань відносна частота події *обов'язково* буде прямувати до ймовірності цієї події. Це прямування також має тільки статистичний характер. Більш точний зв'язок між відносною частотою та ймовірністю події встановлюється теоремою Бернуллі (див. п. 2.41).

1.7. Геометрична ймовірність

Вище відмічалося, що класичне означення ймовірності має, зокрема, той недолік, що це означення не можна застосувати до випробувань з нескінченною кількістю наслідків. З метою подолання цього недоліку вводять геометричні ймовірності, а саме ймовірності потрапляння точки в деяку суцільну множину точок (частину лінії, площини, простору тощо).

Розглянемо деяку плоску фігуру G. І нехай g – частина цієї фігури (рис. 1).

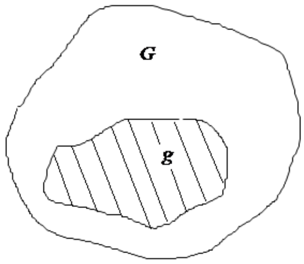


Рис. 1

Нехай у фігуру G випадковим чином потрапляє точка M. Ця точка може опинитися в будь якій точці фігури G. Треба знайти ймовірність того, що точка M потрапить до фігури g. При цьому припускається, що ймовірність потрапляння точки в фігуру прямо пропорційна площі цієї фігури.

Нехай $P(g)$ – ймовірність потрапляння точки M у фігуру g. Тоді $P(g) = k \cdot S(g)$, де $S(g)$ – площа фігури g, а k – коефіцієнт пропорційності. Зокрема, фігура g може співпадати зі всією фігурою G, отже $P(G) = = k \cdot S(G)$. З іншого боку, оскільки у фігуру G точка M за умовою потрапляє обов'язково, то $P(G) = 1$, тобто $k \cdot S(G) = 1$, звідки маємо: $k = 1/S(G)$. Отже,

$$P(g) = k \cdot S(g) = \frac{1}{S(G)} \cdot S(g) = \frac{S(g)}{S(G)}.$$

Якщо фігури G і g розглядаються не на площині, а на лінії, то замість площі використовуються довжина. Якщо фігури розглядаються у тривимірному просторі, то використовується об'єм. У загальному випадку використовується термін *міра*: $mes(g)$. І таким чином отримуємо формулу:

$$P(g) = \frac{mes(g)}{mes(G)}.$$

Ця формула називається формулою *геометричної ймовірності*.

Приклади

1. На відрізку довжиною L розташовано менший відрізок довжиною $l < L$. Знайти ймовірність того, що точка, яку навмання поставлено на більший відрізок, потрапить також на менший.

Згідно з формулою геометричної ймовірності маємо:

$$P = \frac{l}{L}.$$

2. У круг радіусу R поміщено концентричний з ним круг радіусу $r < R$. Знайти ймовірність того, що точка, яку навмання кинуто в більший круг, потрапить в кільце, яке утворено між колами цих кругів (рис. 2).

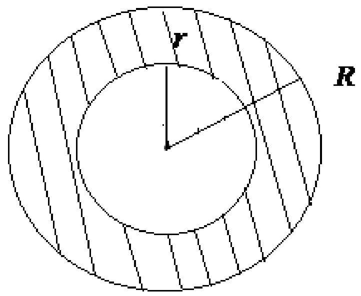


Рис. 2

Згідно з формулою геометричної ймовірності маємо:

$$P = \frac{S_1}{S_2},$$

де S_1 – площа кільця, S_2 – площа великого круга. Маємо: $S_2 = \pi R^2$, $S_1 = \pi R^2 - \pi r^2$, отже

$$P = \frac{\pi R^2 - \pi r^2}{\pi R^2} = 1 - \left(\frac{r}{R}\right)^2.$$

3. У куб вписано кулю. Знайти ймовірність того, що точка, яку навмання кинуто в куб, потрапить також у кулю.

Згідно з формулою геометричної ймовірності маємо:

$$P = \frac{V_1}{V_2},$$

де V_1 – об'єм кулі, V_2 – об'єм куба. Нехай a – довжина ребра куба, тоді радіус кулі $r = a/2$. Отже $V_1 = 4\pi r^3/3 = \pi a^3/6$, $V_2 = a^3$, і таким чином:

$$P = \frac{\pi a^3/6}{a^3} = \frac{\pi}{6}.$$

Ми бачимо, що ця ймовірність не залежить від a, тобто розмірів куба.

4. На площину нанесено сітку квадратів зі стороною a. На цю площину навмання кинуто монету радіусу $r < a/2$. Знайти ймовірність того, що монета не перетинатиме жодної з сторін квадратів.

Розглянемо квадрат зі стороною a (рис. 3).

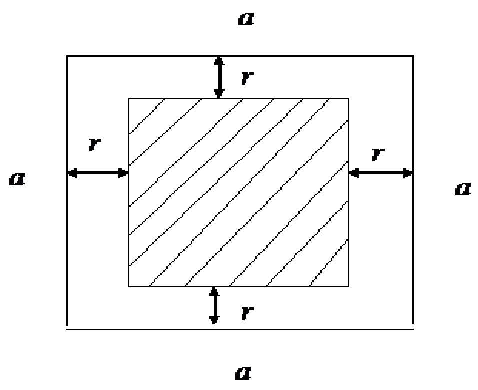


Рис. 3

Для того, щоб кинута в нього монета не перетинала жодної з його сторін, необхідно і достатньо, щоб центр монети потрапив до розташованого в ньому меншого квадрату, сторона якого дорівнює $a-2r$, і ці квадрати мають спільний центр (точку перетину діагоналей). Таким чином, ми приходимо до задачі: знайти ймовірність того, що точка M (центр монети), яка потрапляє до більшого квадрату, потрапить також у менший квадрат.

Маємо: $P = S_1/S_2$, де S_1 – площа меншого квадрату, S_2 – площа більшого квадрату. Отже:

$$P = \frac{(a-2r)^2}{a^2} = \left(1 - \frac{2r}{a}\right)^2.$$

5. Задача Бюффона. Площину розкреслено паралельними прямими, відстань між двома сусідніми дорівнює $2a$. На площину навмання кидають відрізок довжиною $2l$ $(l < a)$. Знайти ймовірність того, що цей відрізок перетинатиме яку-небудь з прямих.

Нехай точка M – середина відрізку, x – відстань від точки M до найближчої прямої, φ – кут, який утворено відрізком з цією прямою (рис. 4).

Положення відрізку, який кинуто, повністю визначається заданням певних значень x і φ, причому x приймає значення від 0 до a, а φ – від 0 до π. Інакше кажучи, точка M може потрапити в будь-яку точку прямокутника зі сторонами a і π. Таким чином, цей прямокутник можна розглядати як фігуру G, точками якої є всі можливі положення точки M. Очевидно, що площа фігури G дорівнює $a\pi$.

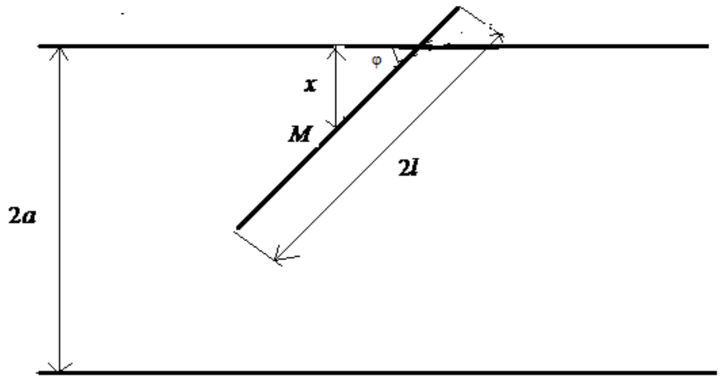


Рис. 4

Знайдемо тепер фігуру g, кожна точка якої сприяє події, що нас цікавить, тобто фігуру, до якої має потрапити точка M для того, щоб відрізок, який кинуто, перетинав пряму лінію. З рис. 4 видно, що відрізок перетинатиме найближчу до нього пряму тоді і тільки тоді, коли $x \le l \sin\varphi$. Таким чином, якщо ввести систему координат φOx, то точка M, потрапляючи до прямокутника $OABC$, має потрапити до криволінійної трапеції, яку обмежено графіком функції $x = l\sin\varphi$ (рис. 5).

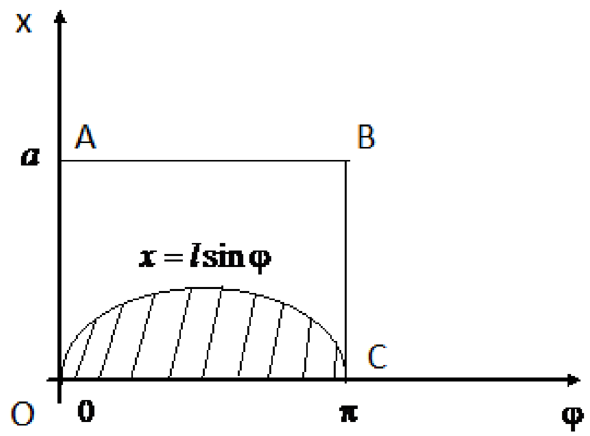


Рис. 5

Саме цю трапецію можна розглядати як фігуру g. Її площа:

$$S(g) = \int_0^{\pi} l\sin\varphi\, d\varphi = -l\cos\varphi\Big|_0^{\pi} = -l(\cos\pi - \cos 0) = -l(-1-1) = 2l.$$

Отже, шукана ймовірність дорівнює:

$$P = \frac{S(g)}{S(G)} = \frac{2l}{a\pi}.$$

6. На відрізок $[0, L]$ навмання поставлено дві точки, координати яких x та y. Знайти ймовірність того, що з відрізків, на які ці точки розбивають відрізок $[0, L]$, можна побудувати трикутника.

Для того, щоб з трьох відрізків можна було побудувати трикутника, необхідно і достатньо, щоб довжина кожного з відрізків була меншою, ніж сума довжин двох інших. Оскільки сума довжин всіх трьох відрізків дорівнює L, то кожен з відрізків за довжиною має бути меншим, ніж $L/2$.

Введемо до розгляду прямокутну систему координат xOy. Координати x та y задовольняють подвійні нерівності:

$0 \le x \le L,\ \ 0 \le y \le L$.

Розглянемо точку $M(x, y)$. Для того, щоб ці нерівності виконувалися, необхідно і достатньо, щоб точка M потрапила до квадрату $OABC$ (рис. 6). Таким чином, цей квадрат можна розглядати як фігуру G, до якої обов'язково потрапляє точка M. Тоді $S(G) = L^2$.

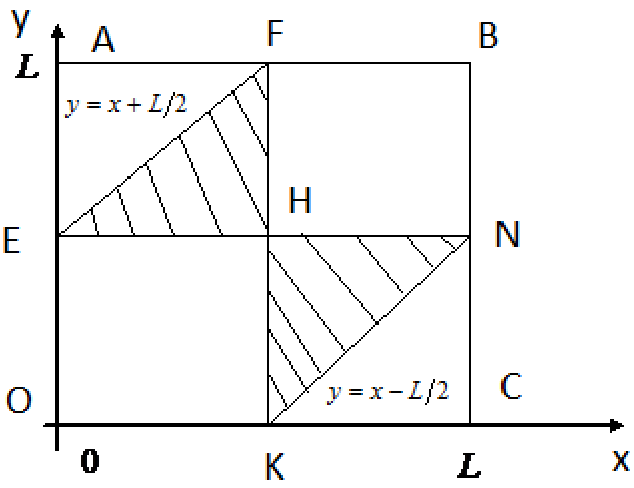
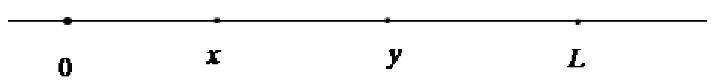


Рис. 6

Розглянемо два можливі випадки.

1) $x < y$:

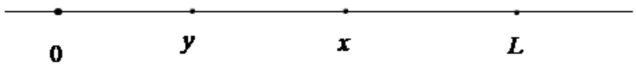


Тоді для здійснення події, що нас цікавить, має бути:

$x < L/2,\ \ y - x < L/2,\ \ L - y < L/2$, або

$$x < L/2,\ y < x + L/2,\ \ y > L/2. \qquad (*)$$

2) $x > y$:



У цьому випадку має бути:

$$y < L/2,\ \ y > x - L/2,\ \ x > L/2. \qquad (**)$$

Нерівності (*) виконано для точок трикутника EFH, а нерівності (**) – для точок трикутника KHN (рис. 6). Таким чином, ці трикутники й є фігура g, до якої має потрапити точка M, щоб відбулася подія, що нас цікавить. Отже:

$$P = \frac{S(g)}{S(G)} = \frac{S_{\Delta EFH} + S_{\Delta KHN}}{S_{\square OABC}} = \frac{1}{4}.$$

7. <u>Задача про зустріч</u>. Два студенти домовилися зустрітися в певному місці між 12 та 13 годинами. Той, що прийшов першим, чекає другого на протязі чверті години, після чого уходить. Знайти ймовірність того, що зустріч відбудеться, якщо кожен студент навмання вибирає мить свого приходу.

Нехай x – мить приходу одного з студентів, y – мить приходу іншого. Оскільки обидва приходи можуть здійснюватися на протязі години, то

$0 \le x \le 1,\ 0 \le y \le 1$.

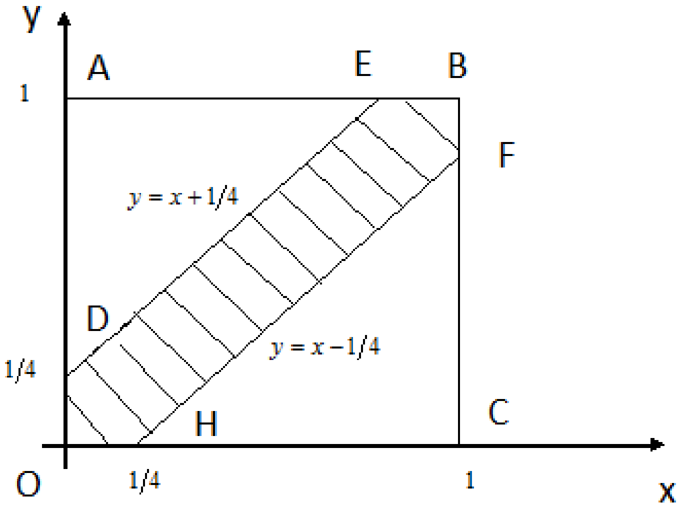


Рис. 7

Зустріч відбудеться тоді і тільки тоді, коли буде виконано нерівність:

$0 \le |x - y| \le \frac{1}{4}$.

Введемо прямокутну систему координат xOy і розглянемо квадрат $OABC$ зі стороною 1 (рис. 7); $S_{OABC} = 1$.

Фігура, яка відповідає останній нерівності, є многокутник $ODEBFH$. Введемо точку $M(x, y)$. Зустріч відбудеться тоді і тільки тоді, коли точка M, потрапляючи до квадрату $OABC$, потрапить при цьому до фігури $ODEBFH$. Отже шукана ймовірність:

$$P = \frac{S_{ODEBFH}}{S_{OABC}} = S_{ODEBFH} = 1 - 2S_{\Delta DAE} = 1 - 2 \cdot \frac{1}{2} \cdot \left(\frac{3}{4}\right)^2 = 1 - \frac{9}{16} = \frac{7}{16}.$$

8. Навмання взято два додатні числа x, y, кожне з яких не перевищує 2. Знайти ймовірність того, що їх добуток xy буде не більше за одиницю, а частка y/x – не більша, ніж 2.

Введемо прямокутну систему координат xOy. За умовою:

$0 < x \le 2,\ \ 0 < y \le 2$.

Подія, що нас цікавить, відбудеться тоді і тільки тоді, коли

$xy \le 1,\ \ \frac{y}{x} \le 2$,

або

$y \le \frac{1}{x},\ \ y \le 2x$.

Розглянемо точку $M(x, y)$. Для здійснення події необхідно і достатньо, щоб ця точка, потрапляючи до квадрату $OABC$, потрапила до фігури $OEHC$ (рис. 8).

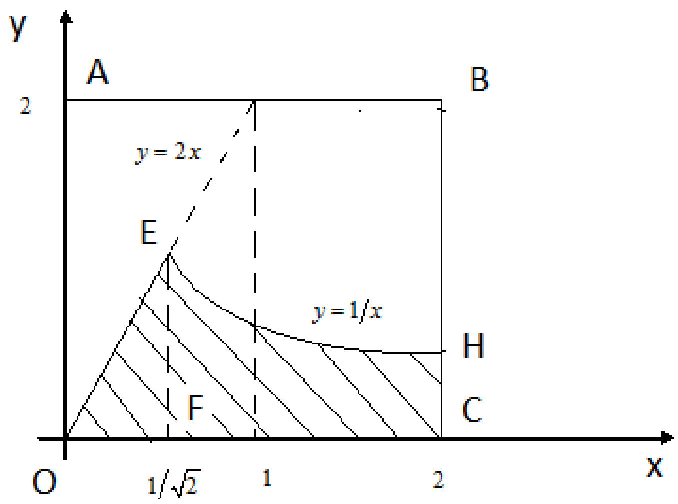


Рис. 8

Отже шукана ймовірність:

$$P=\frac{S_{OEHC}}{S_{OABC}}=\frac{1}{4}\cdot S_{OEHC}=\frac{1}{4}\left(S_{\Delta OEF}+S_{FEHC}\right)=\frac{1}{4}\left(\frac{1}{2}\cdot\frac{1}{\sqrt{2}}\cdot\sqrt{2}+\int\limits_{1/\sqrt{2}}^{2}\frac{dx}{x}\right)=$$

$$=\frac{1}{4}\left(\frac{1}{2}+\ln x\Big|_{1/\sqrt{2}}^{2}\right)=\frac{1}{4}\left(\frac{1}{2}+\ln 2-\ln\frac{1}{\sqrt{2}}\right)=\frac{1}{4}\left(\frac{1}{2}+\frac{3}{2}\ln 2\right)=\frac{1+3\ln 2}{8}\approx 0{,}38\,.$$

Помітимо, що задачі 7, 8 за своїм формулюванням, здавалося б, зовсім не геометричні. Але допускають геометричну інтерпретацію, яка й дає можливість використання формули геометричної ймовірності.

1.8. Умовна ймовірність. Ймовірність добутку подій

Означення. *Умовною ймовірністю події B за умови A* називається ймовірність події B, яку обчислено за умови, що подія A вже відбулася.

Позначається умовна ймовірність: $P(B/A)$, або $P_A(B)$.

Приклад. В урні 10 куль – 6 білих та 4 чорних. Навмання одна за одною з урни виймаються кулі та відкладаються, тобто в урну не повертаються. Нехай подія A_j – j-а куля, що дістали, біла, подія B_j – j-а куля, що дістали, чорна. Знайти наступні ймовірності: $P(A_1)$, $P(B_1)$, $P(A_2/A_1)$, $P(A_2/B_1)$, $P(B_2/A_1)$, $P(B_2/B_1)$, $P(A_3/A_1A_2)$.

Маємо: $P(A_1)=6/10$; $P(B_1)=4/10$; $P(A_2/A_1)=5/9$ (після 1-го виймання залишилося 9 куль, серед них 5 білих); $P(A_2/B_1)=6/9$ (після 1-го виймання залишилося 9 куль, серед них 6 білих); $P(B_2/A_1)=4/9$ (після 1-го виймання залишилося 9 куль, серед них 4 чорних); $P(B_2/B_1)=3/9$ (після 1-го виймання залишилося 9 куль, серед них 3 чорних); $P(A_3/A_1\,A_2)=4/8$ (після 2-го виймання залишилося 8 куль, серед них 4 білих).

Теорема. *Умовна ймовірність події B за умови A дорівнює ймовірності події AB, поділеної на ймовірність події A*:

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

Доведення. Нехай n – число всіх рівноможливих наслідків випробування, m – число наслідків, які сприяють події A, k – число наслідків, які сприяють події B, s – число наслідків, які сприяють події AB. Тоді

$$P(B/A)=\frac{s}{m}=\frac{s/n}{m/n}=\frac{P(AB)}{P(A)},$$

що й треба було довести.

Наслідок. *Ймовірність добутку подій A і B, тобто події AB, дорівнює ймовірності події A, помноженої на умовну ймовірність події B за умови A*:

$$P(AB) = P(A)P(B/A). \qquad (*)$$

Приклад. Студент при підготовці до заліку вивчив 25 питань з 30, що були у програмі. Знайти ймовірність того, що він відповість на 2 поспіль питання, що запропоновано викладачем.

Позначимо: A_1 – подія, що полягає в тому, що студент відповість на перше питання, A_2 – подія, яка полягає в тому, що він відповість на друге питання. Тоді подія, що нас цікавить, це A_1A_2. Маємо:

$$P(A_1A_2) = P(A)P(A_2/A_1) = \frac{25}{30} \cdot \frac{24}{29} = \frac{20}{29} .$$

Зауваження 1. Застосовуючи формулу (*) до події BA, дістанемо:

$$P(BA) = P(B)P(A/B).$$

Але, оскільки $AB = BA$, то $P(AB) = P(B)P(A/B)$, тобто

$$P(A)P(B/A) = P(B)P(A/B). \qquad (**)$$

Зауваження 2. Формула (*) легко узагальнюється на випадок добутку декількох подій. Наприклад, для трьох подій A, B, C маємо:

$$P(ABC) = P(A)P(B/A)P(C/AB).$$

Означення. Подія B називається *незалежною від події A*, якщо поява події A не змінює ймовірності події B, тобто, якщо умовна ймовірність події B за умови A дорівнює безумовній ймовірності події B:

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

Підставляючи це співвідношення до формули (**), дістанемо:

$$P(A)P(B) = P(B)P(A/B),$$

звідки

$$P(A/B) = P(A),$$

а це означає, що й подія A не залежить від події B. Отже, якщо подія B не залежить від події A, то й подія A не залежить від події B. Тобто властивість незалежності події є взаємною.

Для незалежних подій A і B формула (*) набуває вигляду:

$$P(AB) = P(A)P(B), \qquad (***)$$

тобто *ймовірність сумісної появи двох незалежних подій дорівнює добутку їх ймовірностей*.

Рівність (***) приймається також в якості означення незалежних подій.

Означення. Дві події називаються *незалежними*, якщо ймовірність їх сумісної появи дорівнює добутку ймовірностей цих подій; в протилежному випадку події називаються залежними.

Приклад. Два стрілка стріляють по одній цілі по одному разу. Перший стрілок влучає в ціль з ймовірністю 0,8, а другий – з ймовірністю 0,6. Знайти ймовірність того, що обидва стрілки влучать в ціль.

Нехай подія A полягає в тому, що в ціль влучив перший стрілок, подія B – в ціль влучив другий стрілок. Оскільки жодна з цих подій ніяк не впливає на ймовірність іншої (стрілки влучають або не влучають в ціль повністю незалежно один від одного), то події A і B незалежні, отже згідно з (***):

$$P(AB) = P(A)P(B) = 0{,}8 \cdot 0{,}6 = 0{,}48.$$

Означення. Декілька подій $A_1, A_2, ..., A_n$ називаються *попарно незалежними*, якщо кожні два з них незалежні.

Означення. Декілька подій $A_1, A_2, ..., A_n$ називаються *незалежними в сукупності*, якщо вони попарно незалежні, і, крім того, кожна з цих подій є незалежною зі всіма можливими добутками решти цих подій.

Наприклад, якщо події A_1, A_2, A_3 незалежні в сукупності, то незалежними є події: A_1 і A_2, A_1 і A_3, A_2 і A_3, A_1 і A_2A_3, A_2 і A_1A_3, A_3 і A_1A_2.

Якщо події $A_1, A_2, ..., A_n$ незалежні в сукупності, то вони є й попарно незалежними. Але зворотне твердження несправедливе, тобто з попарної незалежності подій не випливає, взагалі кажучи, їх незалежність в сукупності.

Приклад. Нехай в скриньці є 4 кулі – одна червона, одна синя, одна чорна і одна пофарбована в усі три ці кольори. Навмання з скриньки виймається одна куля. Знайдемо ймовірність того, що вона має червоний колір. Нехай подія A – куля має червоний колір, подія B – куля має синій колір, подія C – куля має чорний колір. Оскільки червоний колір мають дві кулі з 4-х, то $P(A) = 1/2$. Аналогічно $P(B) = 1/2$, $P(C) = 1/2$. Припустимо тепер, що вийнята куля має синій колір, тобто подія B вже відбулася. Знайдемо $P(A/B)$. Синій колір мають тільки дві кулі, а з цих двох червоний має тільки одна. Отже $P(A/B) = 1/2$, тобто події A і B незалежні. Аналогічно і події A і C, B і C також незалежні. Таким чином, події A, B, C є попарно незалежними. Чи є вони незалежними в сукупності? Знайдемо $P(A/BC)$. Синій та чорний колір водночас має тільки одна куля, вона ж обов'язково має й червоний колір, отже $P(A/BC) = 1$, тобто $P(A/BC) \neq P(A)$. Отже, події A, B, C не є незалежними в сукупності.

Теорема. *Якщо події $A_1, A_2, ..., A_n$ незалежні в сукупності, то*

$$P(A_1A_2 \cdots A_n) = P(A_1)P(A_2) \cdots P(A_n).$$

Доведення. Розглянемо три події A, B, C. Сумісність подій A, B, C рівносильна сумісності подій AB і C, отже:

$$P(ABC) = P(AB \cdot C).$$

Оскільки події A, B, C незалежні в сукупності, то, зокрема, незалежні події AB і C, а також A і B. За формулою (***) маємо:

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

Для випадку n подій доведення проводиться аналогічно. Зауважимо, що якби події $A_1, A_2, ..., A_n$ були незалежними не в сукупності, а тільки попарно, то ця теорема не була б справедливою. Дійсно, повертаючись до вищеописаного прикладу з скринькою, де 4 кулі 3-х кольорів, маємо:

$$P(A)P(B)P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8},$$

але в той же час

$$P(ABC) = \frac{1}{4},$$

оскільки водночас всі три кольори має тільки одна куля з 4-х.

1.9. Ймовірність суми подій

Розглянемо спочатку питання про ймовірність суми двох подій.

Теорема. *Ймовірність суми двох подій дорівнює сумі їх ймовірностей без ймовірності їх добутку:*

$$P(A + B) = P(A) + P(B) - P(AB).$$

Доведення. Нехай n – загальне число всіх елементарних наслідків випробування, m – число наслідків, які сприяють події A, k – число наслідків, які сприяють події B, s – число наслідків, які сприяють сумісній появі подій A і B, тобто AB. Сума подій A і B полягає в появі хоча б однієї з цих подій. Тому в число наслідків, які сприяють $A + B$, увійдуть всі наслідки, які сприяють події A, і всі наслідки, які сприяють події B, всього таких наслідків $m + k$. Помітимо, що до числа наслідків, які сприяють події A, увійдуть і наслідки, які сприяють події AB, ці ж наслідки увійдуть і до числа наслідків, які сприяють події B. Тому в сумі $m + k$ число наслідків, що сприяють AB, ураховано двічі. Отже одного разу треба позбавитися, і тому число наслідків, що сприяють $A + B$, дорівнює $m + k - s$. Таким чином:

$$P(A + B) = \frac{m + k - s}{n} = \frac{m}{n} + \frac{k}{n} - \frac{s}{n} = P(A) + P(B) - P(AB),$$

що й треба було довести.

Приклад. З колоди в 36 карт навмання виймається одна карта. Знайти ймовірність того, що ця карта або король, або бубна.

Нехай подія A – карта є королем, подія B – карта є бубною. Нас цікавить $P(A+B)$. Маємо: $P(A)=1/9$, $P(B)=1/4$, $P(AB)=1/36$. Отже:

$$P(A+B)=P(A)+P(B)-P(AB)=\frac{1}{9}+\frac{1}{4}-\frac{1}{36}=\frac{12}{36}=\frac{1}{3}.$$

Наслідок 1. *Якщо події A і B несумісні, то ймовірність їх суми дорівнює сумі їх ймовірностей, тобто:*

$$P(A+B)=P(A)+P(B).$$

Дійсно, оскільки A і B несумісні, то $P(AB)=0$, отже з теореми про ймовірність суми маємо $P(A+B)=P(A)+P(B)$.

Приклад. Стрілок стріляє по цілі, яку розділено на 3 частини, що не перетинаються. Ймовірність влучання в першу частину дорівнює 0,45, в другу – 0,35. Знайти ймовірність того, що стрілок при пострілі влучить або в першу, або в другу частину.

Нехай подія A – стрілок влучив в першу частину, подія B – стрілок влучив в другу частину. Оскільки частини не перетинаються, то події A і B несумісні, отже:

$$P(A+B)=P(A)+P(B)=0,45+0,35=0,80.$$

Наслідок 2. *Ймовірність появи однієї з декількох попарно несумісних подій дорівнює сумі їх ймовірностей:*

$$P(A_1+A_2+\ldots+A_n)=P(A_1)+P(A_2)+\ldots+P(A_n).$$

Розглянемо три попарно несумісні події A, B, C. Поява однієї з цих подій еквівалентна появі однієї з двох подій $A+B$ і C, тому згідно з наслідком 1:

$$P(A+B+C)=P((A+B)+C)=P(A+B)+P(C)=P(A)+P(B)+P(C).$$

Для довільного числа попарно несумісних подій доведення проводиться методом математичної індукції.

Наслідок 3. *Сума ймовірностей взаємно протилежних подій дорівнює 1:*

$$P(A)+P(\overline{A})=1.$$

Доведення. Взаємно протилежні події є несумісними, отже $P(A+\overline{A})=$ $=P(A)+P(\overline{A})$. З іншого боку $A+\overline{A}=E$ («теорема Вінні-Пуха»), отже, $P(A+\overline{A})=1$, звідки й випливає твердження.

Таким чином, ймовірність події $\overline{A}$ знаходиться за формулою:

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

Наслідок 4. *Ймовірність появи хоча б однієї з подій* $A_1, A_2, ..., A_n$, *які є незалежними в сукупності, дорівнює різниці між одиницею та добутком ймовірностей протилежних подій* $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$:

$$P(A_1 + A_2 + \ldots + A_n) = 1 - P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_n).$$

Доведення. Нехай $A = A_1 + A_2 + \ldots + A_n$. Подія A та подія $\overline{A}_1\,\overline{A}_2 \cdots \overline{A}$ (жодна з подій $A_1, A_2, ..., A_n$ не відбулася) протилежні, отже:

$$P(A) = 1 - P(\overline{A}_1\,\overline{A}_2 \cdots \overline{A}_n) = 1 - P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_n)$$

(на підставі того, що події $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$ також є незалежними в сукупності, отже, ймовірність їх добутку дорівнює добутку їх ймовірностей), що й треба було довести.

Якщо, зокрема, $P(A_j) = p \;\; (j = 1,2,...,n)$, то $P(\overline{A}_j) = q = 1 - p \;\; (j = 1,2,...,n)$, і тоді

$$P(A) = 1 - q^n.$$

Приклад. Три гармати роблять по одному пострілу. Перша гармата влучає в ціль з ймовірністю 0,8, а для другої та третьої ці ймовірності дорівнюють відповідно 0,7 та 0,9. Знайти ймовірність хоча б одного влучання в ціль при одному залпі з всіх трьох гармат.

Нехай подія A_1 – в ціль влучила перша гармата, подія A_2 – в ціль влучила друга гармата, подія A_3 – в ціль влучила третя гармата. Ці події є незалежними в сукупності. Ймовірності протилежних подій: $P(\overline{A}_1) = 0,2$, $P(\overline{A}_2) = 0,3$, $P(\overline{A}_3) = 0,1$. Тому згідно з наслідком 4:

$$P(A_1 + A_2 + A_3) = 1 - P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3) = 1 - 0,2 \cdot 0,3 \cdot 0,1 = 1 - 0,006 = 0,994.$$

1.10. Задачі на обчислення ймовірності суми та добутку подій

Задача 1. В ящику лежать 10 деталей, з яких 4 стандартні. Навмання беруть 6 деталей. Знайти ймовірність того, що серед цих 6 не більше 2 стандартних.

Нехай A_k – подія, яка полягає в тому, що серед 6 деталей, що взято, рівно k стандартних. Тоді подія, що нас цікавить: $A_0 + A_1 + A_2$. Події A_0, A_1, A_2 несумісні, отже:

$$P(A_0 + A_1 + A_2) == P(A_0) + P(A_1) + P(A_2).$$

Маємо (див. задачу про вибірку, п. 1.5):

$$P(A_0)=\frac{C_4^0\cdot C_6^6}{C_{10}^6}=\frac{1}{\frac{10!}{6!\cdot 4!}}=\frac{1}{210},$$

$$P(A_1)=\frac{C_4^1\cdot C_6^5}{C_{10}^6}=\frac{4\cdot 6}{210}=\frac{24}{210},$$

$$P(A_2)=\frac{C_4^2\cdot C_6^4}{C_{10}^6}=\frac{\frac{4!}{2!\cdot 2!}\cdot\frac{6!}{4!\cdot 2!}}{210}=\frac{90}{210}.$$

Отже:

$$P(A_0+A_1+A_2)=\frac{1}{210}+\frac{24}{210}+\frac{90}{210}=\frac{115}{210}=\frac{23}{42}.$$

Задача 2. В умовах попередньої задачі знайти ймовірність того, що серед 6 деталей, що взято, хоча б одна стандартна.

Очевидно, цього разу нас цікавить подія $A=A_1+A_2+A_3+A_4+A_5+A_6$. І тоді нам треба обчислювати ймовірність суми 6 подій. Але можна зробити інакше. Розглянемо протилежну подію $\overline{A}$. Очевидно, що $\overline{A}=A_0$ (серед 6 деталей, що взято, нема жодної стандартної). Отже:

$$P(A)=1-P(\overline{A})=1-P(A_0)=1-\frac{1}{210}=\frac{209}{210}.$$

Задача 3. Студент вивчив 20 питань з 25, що їх є у програмі заліку. Викладач приймає залік за наступною схемою. Студенту задається питання. Якщо він на нього відповідає правильно, то отримує залік. Якщо правильної відповіді нема, то задається ще одне питання. Якщо на друге питання студент дає правильну відповідь, то отримує залік, а у протилежному випадку – незалік. Знайти ймовірність того, що студент здасть залік.

Нехай подія A – студент відповів на перше питання, подія B – студент відповів на друге питання. Тоді подія, що полягає в тому, що залік буде отримано, еквівалентна події $A+\overline{A}B$. Події A і $\overline{A}B$ несумісні, а події $\overline{A}$ і B залежні, отже:

$$P(A+\overline{A}B)=P(A)+P(\overline{A}B)=P(A)+P(\overline{A})P(B/\overline{A})=\frac{20}{25}+\frac{5}{25}\cdot\frac{20}{24}=\frac{29}{30}.$$

Задача 4. У скриньці є 10 кульок з номерами від 1 до 10. Навмання одна за одною дістаються 3 кульки та назад до скриньки не повертаються. Знайти ймовірність того, що послідовно з'являться кульки з номерами 1, 4, 7.

Нехай подія A – перша кулька має номер 1, подія B – друга кулька має номер 4, подія C – третя кулька має номер 7. Тоді нас цікавить ймовірність події ABC. Маємо:

$$P(ABC)=P(A)P(B/A)P(C/AB)=\frac{1}{10}\cdot\frac{1}{9}\cdot\frac{1}{8}=\frac{1}{720}.$$

Задача 5. Два стрілка стріляють по одному разу в одну ціль. Перший стрілок влучає в ціль з ймовірністю 0,7, а другий – з ймовірністю 0,9. Знайти ймовірність наступних подій:

A – влучили обидва стрілки;

B – влучив тільки перший стрілок;

C – влучив тільки другий стрілок;

D – влучив тільки один стрілок;

F – влучив хоча б один стрілок;

G – жоден стрілок не влучив.

Нехай подія H_1 – влучив перший стрілок, подія H_2 – влучив другий стрілок. За умовою: $P(H_1)=0,7$, $P(H_2)=0,9$. Події H_1 і H_2 сумісні та незалежні. Тому маємо:

$$P(A)=P(H_1H_2)=P(H_1)P(H_2)=0,7\cdot 0,9=0,63;$$

$$P(B)=P\left(H_1\overline{H_2}\right)=P(H_1)P\left(\overline{H_2}\right)=P(H_1)(1-P(H_1))=$$
$$=0,7(1-0,9)=0,7\cdot 0,1=0,07;$$

$$P(C)=P\left(\overline{H_1}H_2\right)=P\left(\overline{H_1}\right)P(H_2)=(1-P(H_1))P(H_2)=$$
$$=(1-0,7)\cdot 0,9=0,3\cdot 0,9=0,27;$$

$$P(D)=P(B+C)=P(B)+P(C)=0,07+0,27=0,34;$$

$$P(F)=P(H_1+H_2)=P(H_1)+P(H_2)-P(H_1H_2)=0,7+0,9-0,63=0,97;$$

$$P(G)=P\left(\overline{H_1}\overline{H_2}\right)=P\left(\overline{H_1}\right)P\left(\overline{H_2}\right)=0,3\cdot 0,1=0,03.$$

Зауважимо, що події F і G є взаємно протилежними, тому $P(G)$ можна було б знайти також за формулою: $P(G)=1-P(F)$.

1.11. Формула повної ймовірності

Розглянемо таку задачу. Нехай для перевірки на якість виріб потрапляє з рівними ймовірностями до одного з двох експертів. А експерти різні за своїми здібностями; і вони по-різному ставляться до своїх обов’язків. Тому ймовірності, що вони визнають виріб якісним, різні. Ймовірність того, що перший експерт визнає виріб якісним, дорівнює 0,4, а відповідна ймовірність для другого експерта – 0,8. Знайти ймовірність того, що виріб буде визнано якісним.

Очевидно, що якби ми наперед знали, до якого саме з двох експертів потрапить виріб, ми б шукану ймовірність одразу ж вказали б – її задано в умові задачі. Але особливістю задачі є як раз те, що ми не знаємо завчасно, який саме експерт буде перевіряти виріб. І ми маємо розв'язувати задачу, інакше кажучи, в умовах невизначеності. Саме таким ситуаціям і присвячено цей параграф.

Введемо наступне поняття.

Означення. Кажуть, що події $B_1, B_2, \ldots, B_n$ утворюють *повну групу*, якщо вони попарно несумісні, але внаслідок випробування хоча б одна з цих подій обов'язково відбудеться.

З означення випливає, що події B_j, B_k $(j \neq k)$ несумісні, і, крім того, $B_1 + B_2 + \ldots + B_n = E$. Отже:

$$P(B_1 + B_2 + \ldots + B_n) = P(B_1) + P(B_2) + \ldots + P(B_n) = 1.$$

Тобто сума ймовірностей подій, що утворюють повну групу, дорівнює одиниці. Фактично внаслідок випробування відбувається не хоча б одна, а точно одна з цих подій. Можна сказати, що події, які утворюють повну групу, це є повний перелік всіх можливих, попарно несумісних наслідків випробування.

Приклад. В ящику 10 куль – 5 білих, 3 чорних та 2 червоних. Навмання дістається одна куля. Події B_1 – ця куля біла, B_2 – ця куля чорна, B_3 – ця куля червона, утворюють повну групу. Дійсно, ці події попарно не сумісні (кожна з куль тільки одного кольору), але внаслідок випробування точно одна з них обов'язково відбудеться (куля, яку дістали, обов'язково буде або білою, або чорною, або червоною). Маємо: $P(B_1) + P(B_2) + P(B_3) = 5/10 + 3/10 + 2/10 = 1$.

Нехай деяка подія A може відбутися при здійсненні однієї з n подій $B_1, B_2, \ldots, B_n$, які утворюють повну групу. Треба знайти ймовірність події A, якщо завчасно невідомо, яка саме з подій $B_1, B_2, \ldots, B_n$ відбулася. Тобто знайти так звану *повну ймовірність* події A.

Назвемо події $B_1, B_2, \ldots, B_n$ *гіпотезами*. За умови подія A може відбутися, якщо здійсниться одна з цих подій. Інакше кажучи, поява події A означає здійснення однієї з n несумісних подій $B_1A, B_2A, \ldots, B_nA$. Маємо:

$$P(A) = P(A(B_1 + B_2 + \ldots + B_n)) = P(AB_1 + AB_2 + \ldots + AB_n) =$$

$$P(AB_1) + P(AB_2) + \ldots +$$

$$+P(AB_n) = P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + \ldots + P(B_n)P(A/B_n).$$

Отже дістали формулу:

$$P(A)=\sum_{k=1}^{n}P(B_k)P(A/B_k),$$

яка називається *формулою повної ймовірності*.

Розглянемо приклади.

Приклад 1. У двох ящиках є кулі. У першому ящику 3 білих та 5 чорних куль, у другому – 6 білих та 7 чорних. З першого ящику навмання дістали одну кулю і переклали в другий ящик, після чого з другого ящика також навмання вийняли одну кулю. Знайти ймовірність того, що ця куля біла.

Введемо гіпотези: B_1 – з першого ящика переклали білу кулю, B_2 – з першого ящика переклали чорну кулю. Знаходимо: $P(B_1)=3/8$, $P(B_2)=5/8$. Нехай подія A – з другого ящика вийнято білу кулю. Тоді $P(A/B_1)=7/14$, $P(A/B_2)=6/14$. Отже за формулою повної ймовірності:

$$P(A)=P(B_1)P(A/B_1)+P(B_2)P(A/B_2)=\frac{3}{8}\cdot\frac{7}{14}+\frac{5}{8}\cdot\frac{6}{14}=\frac{21}{112}+\frac{30}{112}=\frac{51}{112}.$$

Приклад 2. На склад поступає продукція з 3-х заводів. 30% продукції йде з першого заводу, серед цієї продукції 3% браку; 20% продукції йде з другого заводу, серед цієї продукції 5% браку; 50% продукції йде з третього заводу, серед цієї продукції 1% браку. Зі складу навмання береться один виріб. Знайти ймовірність того, що він доброякісний.

Нехай подія A – взятий виріб доброякісний. Введемо гіпотези: B_1 – цей виріб з 1-го заводу, B_2 – цей виріб з 2-го заводу, B_3 – цей виріб з 3-го заводу. За умовою: $P(B_1)=0{,}3$, $P(B_2)=0{,}2$, $P(B_3)=0{,}5$. Далі: $P(A/B_1)=0{,}97$, $P(A/B_2)=0{,}95$, $P(A/B_3)=0{,}99$. Згідно з формулою повної ймовірності:

$$P(A)=P(B_1)P(A/B_1)+P(B_2)P(A/B_2)+P(B_3)P(A/B_3)=$$
$$=0{,}3\cdot 0{,}97+0{,}2\cdot 0{,}95+0{,}5\cdot 0{,}99=0{,}976.$$

1.12. Формули Байєса

Розглянемо таку задачу. Є дві урни з кулями, в кожній урні по 10 куль. В першій 9 білих куль і тільки одна чорна, а в другій навпаки – 1 біла і 9 чорних. Навмання вибирається одна з двох урн. Зрозуміло тоді, що ймовірність вибору кожної урни одна й та ж і дорівнює $1/2$. Тепер припустимо, що з тієї урни, що вибрано, також навмання вибирається одна куля, і ця куля з’являється білою. Тоді питання: з якої урни ймовірніше було взято цю кулю? Зрозуміло, що з першої, адже в ній значно більше можливостей для появи білої кулі. Тобто виходить, що коли ми дістали з урни кулю, і вона з’явилася білою, то ймовірність того, що ми взяли першу урну, не $1/2$, як раніше, а вже інша. Таким чином, якщо A – подія, що з вибраної урни з'явилася біла куля, B_1 – гіпотеза, яка полягає в тому, що було вибрано першу урну, B_2 – гіпотеза, яка полягає в тому, що було вибрано другу урну, то ми можемо стверджувати, що ймовірності $P(B_1), P(B_2)$ гіпотез, які обчислено до здійснення події A, і ймовірності $P(B_1/A), P(B_2/A)$ тих самих гіпотез, які обчислено після здійснення події A – різні. Отже, виникає задача обчислення саме других ймовірностей.

Сформулюємо тепер задачу в загальному випадку. Нехай подія A може відбутися при здійсненні однієї з n гіпотез $B_1, B_2, ..., B_n$. Ймовірності $P(B_1)$, $P(B_2)$, … , $P(B_n)$ цих гіпотез назвемо *апріорними* (від латинського *a priori* – до випробування). Припустимо, що проведено випробування, внаслідок якого відбулася подія A. Треба знайти умовні ймовірності $P(B_1/A)$, $P(B_2/A)$, … , $P(B_n/A)$, які ми назвемо *апостеріорними* (від латинського *a posteriori* – після випробування).

Знайдемо спочатку $P(B_1/A)$. За формулою для ймовірності добутку подій знайдемо:

$$P(AB_1) = P(A)P(B_1/A) = P(B_1)P(A/B_1),$$

звідки

$$P(B_1/A) = \frac{P(B_1)P(A/B_1)}{P(A)} .$$

Обчислюючи тепер $P(A)$ за формулою повної ймовірності, дістанемо:

$$P(B_1/A) = \frac{P(B_1)P(A/B_1)}{P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + ... + P(B_n)P(A/B_n)} .$$

Аналогічно отримуються формули і для інших апостеріорних ймовірностей. Таким чином, маємо:

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{\sum_{k=1}^{n} P(B_k)P(A/B_k)}, \quad i = 1, 2, ..., n.$$

Ці формули називаються *формулами Байєса*[1] .

<u>Приклад</u>. Виріб перевіряється на стандартність одним з двох контролерів. Ймовірність того, що виріб потрапить до 1-го контролера, дорівнює 0,55, а до 2-го – 0,45. Ймовірність того, що стандартний виріб буде визнано стандартним 1-м контролером, дорівнює 0,90, а 2-м – 0,98. Стандартний виріб після перевірки

---

[1] Байєс Томас (1702 – 1761) – англійський математик.

було визнано стандартним. Знайти ймовірність того, що цей виріб перевіряв 2-й контролер.

Нехай подія A – виріб визнано стандартним, гіпотеза B_1 – виріб перевіряв 1-й контролер, гіпотеза B_2 – виріб перевіряв 2-й контролер. Тоді треба знайти $P(B_2/A)$. Згідно з формулами Байєса маємо:

$$P(B_2/A) = \frac{P(B_2)P(A/B_2)}{P(B_1)P(A/B_1) + P(B_i)P(A/B_i)} = \frac{0,45 \cdot 0,98}{0,55 \cdot 0,90 + 0,45 \cdot 0,98} = \frac{0,441}{0,936} \approx 0,47.$$

1.13. Повторення незалежних випробувань. Формула Бернуллі

Припустимо, що ми проводимо серію випробувань, у кожному з яких може відбутися подія A. Причому ймовірність появи події A в кожному випробуванні не залежить від наслідків інших випробувань. Такі випробування називають *незалежними відносно події* A.

Припустимо також, що у кожному з випробувань ймовірність появи події A одна й та ж (хоча, взагалі кажучи, у кожному випробуванні ця ймовірність може бути своя). Позначимо цю ймовірність як p. Тоді ймовірність того, що подія A не відбудеться, у кожному з випробувань теж одна й та ж і дорівнює $q = 1 - p$.

Поставимо задачу: знайти ймовірність того, що при n випробуваннях подія A відбудеться рівно k разів. Позначимо цю ймовірність як $P_n(k)$. Наприклад, $P_5(3)$ означає ймовірність того, що в 5 випробуваннях подія A відбудеться рівно 3 рази.

Розглянемо для спрощення спочатку саме цю ймовірність. Нехай A_j – подія, що полягає в тому, що подія A відбулася у j-му випробуванні. Тоді подію, що нас цікавить, може бути подано у вигляді:

$$\overline{A_1}\,\overline{A_2}A_3A_4A_5 + \overline{A_1}A_2\overline{A_3}A_4A_5 + \overline{A_1}A_2A_3\overline{A_4}A_5 + \overline{A_1}A_2A_3A_4\overline{A_5} + A_1\overline{A_2}\,\overline{A_3}A_4A_5 +$$
$$+A_1\overline{A_2}A_3\overline{A_4}A_5 + A_1\overline{A_2}A_3A_4\overline{A_5} + A_1A_2\overline{A_3}\,\overline{A_4}A_5 + A_1A_2\overline{A_3}A_4\overline{A_5} + A_1A_2A_3\overline{A_4}\,\overline{A_5}.$$

Таким чином, отримали суму 10 подій. Помітимо, що число 10 виникло невипадково – це не що інше, як C_5^3. Оскільки події A_1, A_2, A_3, A_4, A_5 незалежні в сукупності, і $P(A_j) = p$, $P\left(\overline{A_j}\right) = q$, то ймовірність, наприклад, події $\overline{A_1}\,\overline{A_2}A_3A_4A_5$ дорівнює:

$$P\left(\overline{A_1}\,\overline{A_2}A_3A_4A_5\right) = P\left(\overline{A_1}\right)P\left(\overline{A_2}\right)P(A_3)P(A_4)P(A_5) = q \cdot q \cdot p \cdot p \cdot p = p^3q^2.$$

Така ж сама ймовірність і у всієї решти з вищеозначених 10 подій. Ці події є, очевидно, попарно несумісними, тому ймовірність їх суми дорівнює сумі їх ймовірностей, отже:

$$P_5(3) = C_5^3 p^3 q^2 .$$

Аналогічно розглядається і загальний випадок, внаслідок чого отримуємо формулу:

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

Ця формула називається *формулою Бернуллі*[2].

Приклади

1. Стрілок 7 разів стріляє по цілі, причому кожного разу він влучає в ціль з ймовірністю 0,8. Знайти ймовірність того, що він влучить в ціль рівно 5 разів.

Згідно з формулою Бернуллі маємо:

$$P_7(5) = C_7^5 \cdot 0,8^5 \cdot 0,2^2 = \frac{7!}{5! \cdot 2!} \cdot 0,8^5 \cdot 0,2^2 = 21 \cdot 0,8^5 \cdot 0,2^2 \approx 0,275 .$$

2. Дві рівносильні хокейні команди НХЛ проводять серію матчів на Кубок Стенлі. Що більш ймовірно для кожної з них – виграти два матчі з чотирьох, або три матчі з шести? Нічиї, згідно з правилами цих змагань, виключено.

Оскільки команди рівносильні, то кожна з них виграє матч з ймовірністю $p = 1/2$. Отже ймовірність виграти два матчі з чотирьох:

$$P_4(2) = C_4^2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^2 = \frac{4!}{2! \cdot 2!} \cdot \frac{1}{2^4} = \frac{6}{16} .$$

А ймовірність виграти три матчі з шести:

$$P_6(3) = C_6^3 \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^3 = \frac{6!}{3! \cdot 3!} \cdot \frac{1}{2^6} = \frac{5}{16} .$$

Оскільки $P_4(2) > P_6(3)$, то більш ймовірно виграти 2 матчі з 4-х, ніж 3 з 6.

3. Монету кидають 5 разів. Знайти ймовірність того, що «герб» випаде не менш, ніж два рази.

Нехай подія A_k – «герб» випав рівно k разів. Тоді подія, що нас цікавить є $A_2 + A_3 + A_4 + A_5$. А протилежна їй подія – $A_0 + A_1$. Знайдемо ймовірність саме цієї події:

$$P(A_0 + A_1) = P(A_0) + P(A_1) = C_5^0 \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^5 + C_5^1 \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{2^5} + \frac{5}{2^5} = \frac{6}{2^5} .$$

Отже шукана ймовірність:

---

[2] Бернуллі Якоб I (1654 – 1705) – видатний швейцарський математик, один з представників славнозвісної сім'ї Бернуллі.

$$P = 1 - P(A_0 + A_1) = 1 - \frac{6}{2^5} = 1 - \frac{6}{32} = \frac{13}{16}\ .$$

4. На відрізок AB довжиною a навмання кинуто 6 точок. Знайти ймовірність того, що 4 точки опиняться від точки A на відстані менш, ніж x $(x < a)$, а 2 точки – на відстані більшій, ніж x. Припускається, що ймовірність потрапляння точки на відрізок пропорційна довжині цього відрізку і не залежить від його розташування.

Нехай точка C лежить від точки A на відстані x. Тоді вказані 4 точки мають опинитися на відрізку AC (рис. 9).

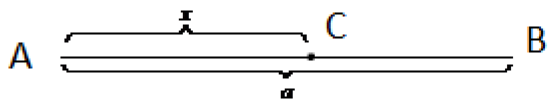


Рис. 9

Знайдемо ймовірність p потрапляння однієї точки на відрізок AC, якщо вона обов'язково потрапляє на відрізок AB. Згідно з формулою геометричної ймовірності:

$$p = \frac{x}{a}.$$

Тепер згідно з формулою Бернуллі маємо:

$$P_6(4) = C_6^4 \cdot \left(\frac{x}{a}\right)^4 \left(1 - \frac{x}{a}\right)^2 = \frac{6!}{4! \cdot 2!} \cdot \left(\frac{x}{a}\right)^4 \left(\frac{a-x}{a}\right)^2 = 15 \cdot \frac{x^4 (a-x)^2}{a^6}\ .$$

Зокрема, якщо, наприклад, $a = 10,\ x = 7$, то

$$P_6(4) = 15 \cdot \frac{7^4 \cdot 3^2}{10^6} \approx 0,324.$$

5. Задача Банаха[3]. Математик, що палить, носить з собою 2 коробки сірників. Кожного разу, коли він хоче дістати сірник, він навмання дістає одну з коробок. Знайти ймовірність того, що, коли він першого разу витягне порожню коробку, у другій коробці залишиться рівно r сірників, якщо початково у кожній коробці було по n сірників $(r \le n)$.

Позначимо як A подію, яка полягає в тому, що сірник дістається з тієї коробки, яка має залишитися порожньою. Ймовірність цієї події в одному випробуванні: $p = 1/2$. Для того, щоб коробка залишилася порожньою, подія A має відбутися n разів. А оскільки у другій коробці при цьому залишилося r

[3] Банах Стефан (1892 – 1945) – видатний польський математик.

сірників, то це означає, що відбулося рівно $2n-r$ діставань сірника. Отже нас цікавить $P_{2n-r}(n)$. Згідно з формулою Бернуллі:

$$P_{2n-r}(n)=C_{2n-r}^{n}\left(\frac{1}{2}\right)^{n}\left(\frac{1}{2}\right)^{2n-r-n}=\frac{C_{2n-r}^{n}}{2^{2n-r}}\ .$$

Якщо, зокрема, $r=n$, то ця ймовірність дорівнює $1/2^{n}$.

1.14. Повторення незалежних випробувань. Формула Пуассона

Розглянемо таку задачу: знайти ймовірність того, що подія A відбудеться 40 разів у 100 випробуваннях, якщо ймовірність появи події A у кожному випробуванні $p=0{,}1$.

Згідно з формулою Бернуллі матимемо:

$$P_{100}(40)=C_{100}^{40}\cdot 0{,}1^{40}\cdot 0{,}9^{60}=\frac{100!}{40!\cdot 60!}\cdot 0{,}1^{40}\cdot 0{,}9^{60}.$$

Звідси видно, що ми змушені мати справу з величезними числами, практично провести такі обчислення дуже складно, до того ж у процесі обчислення накопичуються похибки, внаслідок яких отриманий результат може значно відрізнятися від правильного.

Таким чином, при великих значеннях n формулою Бернуллі користуватися практично неможливо, і у таких випадках користуються іншими формулами. Однією з таких формул є *формула Пуассона*[4].

Припустимо, що ймовірність p дуже мала, $p<<1$. І крім того, добуток np зберігає постійне значення λ: $np=\lambda$. Знайдемо ймовірність $P_n(k)$. Оскільки $p<<1$, то це буде означати, що $k<<n$, тобто подія A трапляється дуже рідко. Скористаємось спочатку формулою Бернуллі:

$$P_n(k)=C_n^k p^k(1-p)^{n-k}=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}=$$

$$=\frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!}p^k(1-p)^{n-k}.$$

Оскільки $np=\lambda$, то $p=\lambda/n$. Отже:

$$P_n(k)=\frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!}\left(\frac{\lambda}{n}\right)^{k}\left(1-\frac{\lambda}{n}\right)^{n-k}.$$

[4] Пуассон Сімеон Дені (1781–1840) – французький математик.

Враховуючи, що n приймає дуже великі значення, замість $P_n(k)$ знайдемо $\lim\limits_{n\to\infty} P_n(k)$.

$$P_n(k) \approx \lim_{n\to\infty} P_n(k) = \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!} \cdot \frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!}\lim_{n\to\infty}\left[1\cdot\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k}\right] =$$

$$= \frac{\lambda^k}{k!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Таким чином:

$$P_n(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}.$$

Це й є формула Пуассона.

Приклади

1. Станок-автомат штампує деталі. Ймовірність того, що деталь буде з дефектом, дорівнює 0,01. Знайти ймовірність того, що серед 200 деталей буде рівно 4 з дефектом.

Маємо: $n = 200,\ p = 0{,}01,\ k = 4,\ \lambda = np = 200\cdot 0{,}01 = 2$.

$$P_{200}(4) \approx \frac{\lambda^k}{k!}e^{-\lambda} = \frac{2^4}{4!}e^{-2} \approx 0{,}09.$$

2. Здійснюється залп з 100 гармат. Ймовірність влучання в ціль для кожної з гармат дорівнює 0,03. Знайти ймовірність того, що буде не менш, ніж 2 влучання в ціль.

Перейдемо до ймовірності протилежної події – вона полягає в тому, що відбудеться менше 2 влучень, тобто або жодного влучання (подія A_0), або одне влучання (подія A_1).

Маємо: $n = 100,\ p = 0{,}03,\ \lambda = 3$.

$$P(A_0) = P_{100}(0) \approx \frac{3^0}{0!}\cdot e^{-3} = e^{-3},\quad P(A_1) = P_{100}(1) \approx \frac{3^1}{1!}\cdot e^{-3} = 3e^{-3}.$$

Події A_0 і A_1 несумісні, отже:

$$P(A_0 + A_1) = P(A_0) + P(A_1) \approx e^{-3} + 3e^{-3} = 4e^{-3}.$$

Таким чином, шукана ймовірність:

$$P(A) \approx 1 - 4e^{-3} \approx 0{,}801.$$

1.15. Повторення незалежних випробувань
Локальна теорема Муавра-Лапласа

Як відмічалося, формула Пуассона використовується при великих значеннях n за умов, що $p \ll 1$, $k \ll n$. Якщо ці умови не виконано, то для обчислення ймовірності $P_n(k)$ при великих n користуються іншою наближеною формулою, яку ми наводимо без доведення:

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{k-np}{\sqrt{npq}}\right), \qquad (*)$$

де функція

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Дослідимо деякі важливі властивості цієї функції. Ця функція, очевидно, додатна: $\varphi(x) > 0$; парна: $\varphi(-x) = \varphi(x)$. Далі $\lim\limits_{x \to \infty} \varphi(x) = 0$, причому прямування до нуля значень функції відбувається досить швидко – на практиці вже при $x > 4$ значення функції можна вважати нулем.

Знайдемо проміжки зростання та спадання та точки екстремуму функції $\varphi(x)$:

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

Звідси видно, що $\varphi'(x) = 0$ при $x = 0$, $\varphi'(x) > 0$ при $x < 0$, $\varphi'(x) < 0$ при $x > 0$. Тобто при $x < 0$ функція зростає, при $x > 0$ функція спадає, і в точці $x = 0$ має максимум, який дорівнює $1/\sqrt{2\pi} \approx 0,3989$.

Знайдемо проміжки опуклості та вгнутості та точки перегину функції $\varphi(x)$:

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

Звідси видно, що $\varphi''(x) = 0$ при $x = \pm 1$, $\varphi''(x) > 0$ при $|x| > 1$, $\varphi''(x) < 0$ при $|x| < 1$. Тобто проміжки $(-\infty, -1), (1, +\infty)$ є проміжками вгнутості, а проміжок $(-1, 1)$ є проміжком опуклості функції $\varphi(x)$. Схематичний графік функції $\varphi(x)$ наведено на рис. 10.

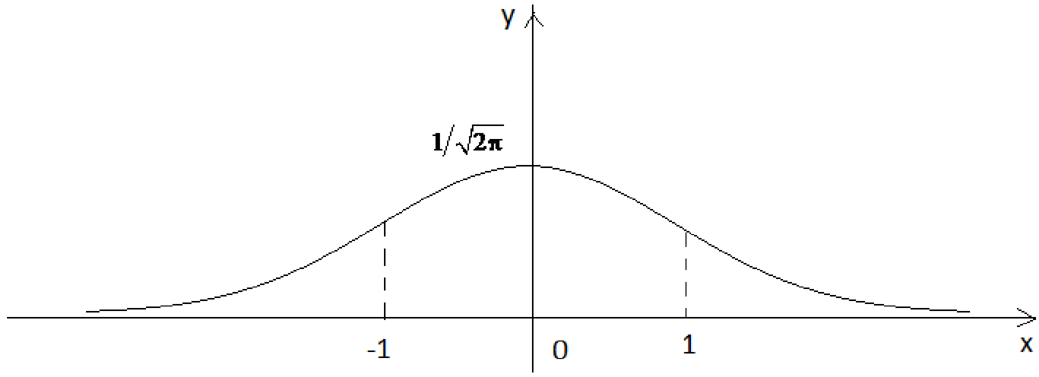


Рис. 10

Функцію $\varphi(x)$ затабульовано, тобто складено детальні таблиці її значень. На практиці для знаходження значень функції користуються саме цими таблицями. Таку таблицю наведено у Додатку (Таблиця 1). В ній наведено значення функції $\varphi(x)$ для сітки значень x від 0 до 4. При $x<0$ користуються властивістю парності цієї функції, а при $x>4$ покладається $\varphi(x)=0$.

Формула (*) називається *локальною теоремою Муавра-Лапласа*[5], а функція $\varphi(x)$ – *диференціальною функцією Лапласа*.

Приклади

1. Ймовірність влучення в ціль при кожному пострілі дорівнює 0,8. Знайти ймовірність того, що при 100 пострілах буде рівно 75 влучень.

За умовою $n=100$, $p=0,8,\ q=0,2,\ k=75$. Помітимо, що тут n велике, а умови $p<<1,\ k<<n$ не виконано, тому формули Бернуллі та Пуассона ми не можемо використати. Скористаємось локальною теоремою Муавра-Лапласа. Маємо:

$$\frac{k-np}{\sqrt{npq}}=\frac{75-100\cdot 0,8}{\sqrt{100\cdot 0,8\cdot 0,2}}=\frac{-5}{\sqrt{16}}=-\frac{5}{4}=-1,25.$$

$$P_{100}(75)\approx\frac{1}{4}\varphi(-1,25)=\frac{1}{4}\cdot 0,1826=0,04565.$$

2. Монета кидається 4 рази. Знайти ймовірність того, що рівно 2 рази з'явиться «герб».

Хоча тут n не є великим ($n=4$), спробуємо все ж скористатися локальною теоремою Муавра-Лапласа. Маємо: $n=4,\ k=2,\ p=q=0,5$,

$$\frac{k-np}{\sqrt{npq}}=\frac{2-4\cdot 0,5}{\sqrt{4\cdot 0,5\cdot 0,5}}=\frac{0}{1}=0,$$

[5] Муавр Абрахам де (1667 – 1754) – англійський математик французького походження. Лаплас П'єр Сімон (1749 – 1827) – видатний французький математик, фізик, астроном і філософ.

$$P_4(2)=\frac{1}{\sqrt{4\cdot 0,5\cdot 0,5}}\varphi(0)=\varphi(0)=0,3989.$$

А тепер розв'яжемо цю задачу за формулою Бернуллі:

$$P_4(2)=C_4^2\cdot 0,5^2\cdot 0,5^2=\frac{4!}{2!\cdot 2!}\cdot\frac{1}{2^4}=\frac{3}{8}=0,375.$$

Ми бачимо значну розбіжність в результатах. Це відбулося внаслідок того, що n не є досить великим числом, тому локальна теорема Муавра-Лапласа дає помітну похибку.

1.16. Повторення незалежних випробувань Інтегральна теорема Муавра-Лапласа

Знову припустимо, що проводиться n випробувань (n досить велике), у кожному з яких подія A з'являється з ймовірністю p. Треба знайти ймовірність $P_n(k_1,k_2)$ того, що число появ події A буде у межах від k_1 до k_2 ($0\le k_1<k_2\le n$). Відповідь на це питання дається наближеною формулою, яку ми наводимо без доведення:

$$P_n(k_1,k_2)\approx\Phi\left(\frac{k_2-np}{\sqrt{npq}}\right)-\Phi\left(\frac{k_1-np}{\sqrt{npq}}\right),$$

(*) де

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt.$$

Формула (*) називається *інтегральною теоремою Муавра-Лапласа*, а функція $\Phi(x)$ – *інтегральною функцією Лапласа* (це інтеграл зі змінною верхньою межею від функції $\varphi(x)$). Дослідимо деякі важливі властивості цієї функції.

1.

$$\Phi(0)=\frac{1}{\sqrt{2\pi}}\int_0^0 e^{-\frac{t^2}{2}}dt=0.$$

2.

$$\Phi(-x)=\frac{1}{\sqrt{2\pi}}\int_0^{-x} e^{-\frac{t^2}{2}}dt\left[\begin{array}{c}\tau=-t,\ dt=-d\tau\\ t=0\Rightarrow\tau=0,\ t=-x\Rightarrow\tau=x\end{array}\right]=\frac{1}{\sqrt{2\pi}}\int_0^{x} e^{-\frac{\tau^2}{2}}(-d\tau)=$$

$$= -\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{\tau^2}{2}}d\tau = -\Phi(x),$$

тобто функція $\Phi(x)$ непарна.

3. Розглянемо:

$$\lim_{x\to+\infty}\Phi(x) = \lim_{x\to+\infty}\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}}\int_0^{+\infty} e^{-\frac{t^2}{2}}dt.$$

Останній невласний інтеграл з нескінченною верхньою межею називається *інтегралом Пуассона*. Можна показати, що його точне значення 0,5. Отже:

$$\lim_{x\to+\infty}\Phi(x) = 0{,}5.$$

Прямування до граничного значення відбувається досить швидко. Практично при $x > 5$ можна вважати, що $\Phi(x)$ дорівнює 0,5.

4.

$$\Phi'(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} = \varphi(x) > 0,$$

отже, функція $\Phi(x)$ є зростаючою на всій числовій прямій.

5.

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

отже, $\Phi''(x) = 0$ при $x = 0$, $\Phi''(x) < 0$ при $x > 0$, $\Phi''(x) > 0$ при $x < 0$. Тобто функція $\Phi(x)$ є вгнутою при $x < 0$ і опуклою при $x > 0$. Точка $x = 0$ є точкою перегину даної функції. Схематичний графік функції $\Phi(x)$ наведено на рис. 11.

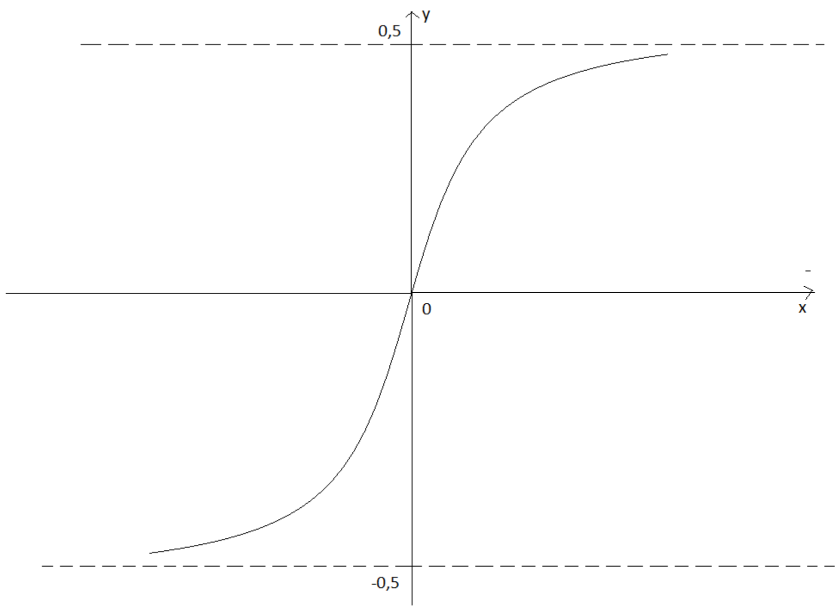


Рис. 11

Зауважимо, що первісна від функції $\exp(-x^2/2)$ не виражається в елементарних функціях. Тому для обчислення значень функції $\Phi(x)$ ми змушені користуватися наближеними методами обчислення інтегралів. За допомогою таких методів отримано досить детальні таблиці наближених значень функції $\Phi(x)$ (див. Додаток, Таблиця 2). Таблиці складено для сітки значень x у межах від 0 до 5. При $x<0$ користуємось властивістю непарності функції $\Phi(x)$, а при $x>5$ покладаємо $\Phi(x)=0,5$.

Приклади

1. Досліджують 500 проб руди. Ймовірність промислового вмісту заліза у кожній пробі дорівнює 0,7. Знайти ймовірність того, що кількість проб з промисловим вмістом заліза буде від 300 до 370.

Скористаємось інтегральною теоремою Муавра-Лапласа. Маємо: $n=500$, $p=0,7$, $q=0,3$, $k_1=300$, $k_2=370$.

$$P_{500}(300,370)\approx\Phi\left(\frac{370-500\cdot 0,7}{\sqrt{500\cdot 0,7\cdot 0,3}}\right)-\Phi\left(\frac{300-500\cdot 0,7}{\sqrt{500\cdot 0,7\cdot 0,3}}\right)=$$

$$=\Phi\left(\frac{20}{\sqrt{105}}\right)-\Phi\left(-\frac{50}{\sqrt{105}}\right)=\Phi(1,95)+\Phi(4,88)=0,4744+0,5=0,9744.$$

2. Ймовірність присутності студента на лекції дорівнює 0,8. Знайти ймовірність того, що з 100 студентів курсу на лекції буде присутнє не менш, ніж 75.

Маємо: $n=100$, $p=0,8$, $q=0,2$, $k_1=75$, $k_2=100$.

$$P_{100}(75,100)=\Phi\left(\frac{100-100\cdot 0,8}{\sqrt{100\cdot 0,8\cdot 0,2}}\right)-\Phi\left(\frac{75-100\cdot 0,8}{\sqrt{100\cdot 0,8\cdot 0,2}}\right)=\Phi(5)-\Phi(-1,25)=$$

$$=\Phi(5)+\Phi(1,25)=0,5+0,3944=0,8944.$$

3. Ймовірність появи додатного результату в кожному з n випробувань дорівнює 0,9. Скільки треба здійснити випробувань, щоб з ймовірністю 0,98 можна було б чекати, що не менш ніж 150 випробувань матимуть додатний результат?

Маємо: $p=0,9$, $q=0,1$, $k_1=150$, $k_2=n$, $P_n(150,n)=0,98$. Скористаємось інтегральною теоремою Муавра-Лапласа:

$$P_n(150,n)=\Phi\left(\frac{n-n\cdot 0,9}{\sqrt{n\cdot 0,9\cdot 0,1}}\right)-\Phi\left(\frac{150-n\cdot 0,9}{\sqrt{n\cdot 0,9\cdot 0,1}}\right)=0,98.$$

Або

$$\Phi\left(\frac{\sqrt{n}}{3}\right)-\Phi\left(\frac{150-0,9n}{0,3\sqrt{n}}\right)=0,98.$$

Очевидно, що $n \ge 150$, отже $\sqrt{n}\big/3 \ge \sqrt{150}\big/3 \approx 4,08$, тому можна покласти $\Phi\left(\sqrt{n}\big/3\right)=0,5$. Отже

$$0,5-\Phi\left(\frac{150-0,9n}{0,3\sqrt{n}}\right)=0,98,$$

звідки

$$\Phi\left(\frac{150-0,9n}{0,3\sqrt{n}}\right)=-0,48.$$

За Таблицею 2 (див. Додаток) знайдемо, що значенню 0,48 функції Лапласа відповідає значення аргументу 2,06, тобто $\Phi(2,06)\approx 0,48$. Отже, враховуючи непарність функції Лапласа, отримуємо:

$$\frac{150-0,9n}{0,3\sqrt{n}}=-2,06,$$

звідки

$$0,9n-0,618\sqrt{n}-150=0.$$

Покладаючи $s=\sqrt{n}$, отримуємо квадратне рівняння:

$$0,9s^2-0,618s-150=0.$$

Додатний корінь цього рівняння $s \approx 13,2578$. Звідси $n \approx 175,77$. Таким чином, шукане число випробувань $n=176$.

1.17. Послідовність незалежних випробувань
Найбільш ймовірне число появ події

Нехай, як і раніше, проводиться серія з n випробувань, у кожному з яких подія A з'являється з ймовірністю p.

Означення. Число k_0 $(0 \le k_0 \le n)$ появ події A називається *найбільш ймовірним*, якщо ймовірність того, що подія A з'явиться в n випробуваннях k_0 разів, не менш, ніж ймовірність того, що подія A з'явиться в n випробуваннях k разів, де k – будь-яке інше можливе число появ події A в n випробуваннях.

Тобто:

$$P_n(k_0)=\max_{0\le k\le n} P_n(k).$$

Можна довести, що число k_0 визначається з подвійної нерівності:

$$np - q \le k_0 \le np + p,$$

причому:

а) якщо число $np - q$ – дробове, то існує тільки одне найбільш ймовірне число k_0;

б) якщо число $np - q$ – ціле, то існують два найбільш ймовірні числа k_0 і $k_0 + 1$;

в) якщо число np – ціле, то найбільш ймовірне число $k_0 = np$.

Приклади

1. Випробується кожен з 15 елементів деякого приладу. Ймовірність того, що елемент витримає випробування, дорівнює 0,9. Знайти найбільш ймовірне число елементів, які витримають випробування.

За умовою $n = 15$, $p = 0,9$, $q = 0,1$. Знайдемо число k_0 з подвійної нерівності:

$$np - q \le k_0 \le np + p.$$

Тобто

$$15 \cdot 0,9 - 0,1 \le k_0 \le 15 \cdot 0,9 + 0,9,$$

або

$$13,4 \le k_0 \le 14,4.$$

Оскільки k_0 – ціле, то $k_0 = 14$.

2. Геолог досліджує 24 зразки мінералів. Ймовірність того, що зразок містить певний хімічний елемент, дорівнює 0,6. Знайти найбільш ймовірне число зразків, які містять цей елемент.

Маємо: $n = 24,\ p = 0,6,\ q = 0,4$. Знайдемо число k_0 з подвійної нерівності $np - q \le k_0 \le np + p$, тобто

$$24 \cdot 0,6 - 0,4 \le k_0 \le 24 \cdot 0,6 + 0,6,$$

або

$$14 \le k_0 < 15.$$

Оскільки $np - q = 14$ – ціле, то найбільш ймовірних чисел два: $k_0 = 14$ і $k_0 + 1 = 15$.

3. Два стрілки водночас стріляють по одній цілі. Ймовірність влучання в ціль при одному пострілі для першого стрілка дорівнює 0,8, а для другого – 0,6. Знайти найбільш ймовірне число залпів, при яких обидва стрілки влучають в ціль, якщо відбудеться 25 залпів.

Влучання 1-го та 2-го стрілка в ціль є незалежні події, тому за теоремою про ймовірність добутку незалежних подій: $p = 0,8 \cdot 0,6 = 0,48$. Таким чином, маємо: $n = 25,\ p = 0,48,\ q = 0,52$. Оскільки $np = 25 \cdot 0,48 = 12$ – ціле, то $k_0 = 12$.

4. Ймовірність появи події у кожному з незалежних випробувань дорівнює 0,3. Знайти n – число незалежних випробувань, при якому найбільш ймовірне число появ події буде 30.

Маємо: $p = 0,3$, $q = 0,7$; $np - q \le 30 \le np + p$. Тобто:

$$0,3n - 0,7 \le 30, \quad 0,3n + 0,3 > 30.$$

З першої нерівності знайдемо:

$$n \le \frac{30,7}{0,3} = \frac{307}{3},$$

отже, оскільки n ціле: $n \le 102$. З другої нерівності:

$$0,3n > 29,7; \quad n > \frac{29,7}{0,3} = \frac{297}{3} = 99.$$

Таким чином: $100 \le n \le 102$.

5. Чому дорівнює ймовірність p появи події у кожному з 39 незалежних випробувань, якщо найбільш ймовірне число появ події в цих випробуваннях дорівнює 25?

Маємо: $n = 39$, $k_0 = 25$. Користуємось подвійною нерівністю:

$$np - q \le k_0 \le np + p,$$

тобто

$$39p - (1 - p) \le 25 \le 39p + p.$$

Або

$$40p - 1 \le 25, \quad 40p > 25.$$

Отже

$$25 < 40p \le 26,$$

тобто:

$$\frac{25}{40} < p \le \frac{26}{40}.$$

Таким чином, шукана ймовірність має задовольняти подвійну нерівність:

$$0,625 < p \le 0,65.$$

6. Прилад складається з 5 незалежно працюючих елементів. Ймовірність відмови елемента в момент включення приладу дорівнює 0,2. Знайти: а) найбільш ймовірне число елементів, що відмовили; б) ймовірність найбільш ймовірного числа елементів, що відмовили.

а) Знайдемо найбільш ймовірне число елементів, що відмовили. Маємо: $n = 5$, $p = 0,2$, $q = 0,8$. Оскільки $np = 5 \cdot 0,2 = 1$ – ціле, то $k_0 = 1$.

б) Знайдемо $P_5(1)$ за формулою Бернуллі:

$$P_5(1) = C_5^1 p^1 q^4 = 5 \cdot 0,2 \cdot 0,8^4 = 0,4096.$$

Контрольні питання до Розділу І

1. Що таке випадкова подія?
2. Які події називаються достовірними, неможливими, протилежними, несумісними?
3. Що таке сума та добуток подій?
4. У чому полягає класичне означення ймовірності події?
5. У чому полягає статистичне означення ймовірності події?
6. Що таке геометрична ймовірність?
7. Що таке умовна ймовірність події?
8. Які дві події називаються незалежними?
9. Які події називаються незалежними попарно?
10. Які події називаються незалежними у сукупності? Чи випливає з незалежності подій попарно їх незалежність у сукупності?
11. Чи завжди ймовірність добутку подій дорівнює добутку їх ймовірностей?
12. Чи завжди ймовірність суми подій дорівнює сумі їх ймовірностей?
13. Що таке повна ймовірність події за наявності гіпотез? За якою формулою її можна обчислити?
14. Що таке апріорні та апостеріорні ймовірності гіпотез? За якими формулами можна обчислити апостеріорні ймовірності?
15. Як обчислити ймовірність того, що подія з’явиться рівно k в n незалежних випробуваннях, якщо ймовірність появи події в кожному випробуванні дорівнює p?
16. Чи доцільно для знаходження цієї ймовірності використовувати формулу Бернуллі, якщо $n = 100$?

Вправи для самостійного розв'язування

1. З слова НАВМАННЯ вибирається навмання одна буква. Яка ймовірність того, що це буква А?

2. Кидаються дві гральні кістки. Знайти ймовірність того, що сума очок, що випали на кістках, ділиться на 3.

3. Дитина грає з буквами А, А, М, М розрізної азбуки. Яка ймовірність того, що, викладаючи їх у випадковому порядку, він отримає слово МАМА?

4. Десять книжок ставляться навмання на одній полиці. Знайти ймовірність того, що при цьому певні три книжки опиняться поряд.

5. В партії з 20 виробів 5 з браком. Знайти ймовірність того, що з взятих навмання 8 виробів 2 будуть з браком.

6. До ліфту 7-поверхового будинку заходять три особи. Кожна з них з однаковою ймовірністю може вийти на будь-якому поверсі, починаючи з другого. Знайти ймовірність того, що всі три особи вийдуть на різних поверхах.

7. В круг навмання кидається точка. Знайти ймовірність того, що ця точка потрапить у вписаний в цей круг правильний трикутник.

8. Штучний супутник Землі рухається вздовж орбіти, яка міститься між 60° північної та 60° південної широти. Вважаючи падіння супутника у будь-яку точку поверхні між вказаними паралелями рівно можливим, знайти ймовірність того, що він впаде вище 30° північної широти.

9. У квадрат з вершинами у точках (0,0), (0,1), (1,0), (1,1) навмання кинуто точку M. Нехай (p,q) – її координати. Знайти ймовірність того, що корені рівняння $x^2 + px + q = 0$ дійсні.

10. В скриньці 12 червоних, 8 зелених та 10 синіх кульок. Навмання вибираються дві кулі. Яка ймовірність того, що ці дві кулі різного кольору, якщо відомо, що не вийнято синю кулю?

11. Відділ технічного контролю перевіряє вироби на стандартність. Ймовірність того, що виріб стандартний, дорівнює 0,9. Знайти ймовірність того, що з двох перевірених виробів тільки один стандартний.

12. Кинуто три гральні кістки. Знайти ймовірність того, що на всіх гранях, що випало, з'явиться одне й те ж число очок.

13. Скільки разів треба кинути гральну кістку, щоб з ймовірністю не менше 0,9 хоча б один раз випало 6 очок.

14. Є дві урни. В першій 3 білих та 2 чорних кулі, в другій 4 білих та 4 чорних. З першої урни до другої, не дивлячись, перекладають дві кулі. Після цього з другої урни беруть одну кулю. Знайти ймовірність того, що ця куля біла.

15. Студент знає не всі екзаменаційні білети. В якому випадку ймовірність витягнути невідомий білет буде для нього менше – коли він тягне білет першим, чи останнім?

16. У магазин надійшло 30% телевізорів із 1-го заводу, серед яких 20% бракованих; 20% – із 2-го заводу, серед яких 10% бракованих; 50% – із 3-го заводу, який має лише 5% браку. Яка ймовірність купити справний телевізор недосвідченим покупцем?

17. До спеціалізованої лікарні потрапляють в середньому 50% хворих із хворобою K, 30% – із хворобою L, 20% – із хворобою M. Ймовірність повного вилікування для цих хвороб складає відповідно 0,7, 0,8, 0,9. Хворий, який потрапив до лікарні, після лікування одужав. Знайти ймовірність того, що у нього була хвороба K.

18. Два стрілка незалежно один від одного стріляють по одній цілі, роблячи кожний по одному пострілу. Ймовірність влучання в ціль для 1-го стрілка дорівнює 0,8, а для другого – 0,4. Після стрільби в цілі було зафіксовано одне влучання. Знайти ймовірність того, що влучив перший стрілок.

19. Спостеріганнями встановлено, що в деякій місцевості у вересні в середньому буває 12 днів, коли йде дощ. Яка ймовірність того, що з випадково взятих у вересні 8 днів 3 дні будуть з дощем?

20. Ймовірність влучання в ціль при одному пострілі дорівнює 0,4. Зроблено 6 незалежних пострілів. Знайти ймовірність хоча б одного влучання в ціль.

21. Ймовірність появи події A хоча б один раз в 5 незалежних випробуваннях дорівнює 0,9. Яка ймовірність появи події A в одному випробуванні, якщо вона однакова в кожному випробуванні?

22. Ймовірність того, що будь-який абонент зателефонує на комутатор на протязі години, дорівнює 0,01. Телефонна станція обслуговує 800 абонентів. Яка ймовірність того, що на протязі години зателефонують 5 абонентів?

23. Ймовірність виробництва бракованого виробу дорівнює 0,008. Знайти ймовірність того, що серед 1000 виробів бракованих буде не більше, ніж 3.

24. До перших класів має бути прийнято 200 дітей. Знайти ймовірність того, що серед них буде рівно 100 дівчат, якщо ймовірність народження хлопчика 0,515.

25. Яка ймовірність того, що у стовпчику із 100 навмання відібраних монет, число монет, розташованих «гербом» вверх, буде від 45 до 55?

26. Гральну кістку кидають 80 разів. Знайти наближено межі, в яких число появ «шестірки» буде міститись з ймовірністю 0,9973.

Розділ ІІ. Випадкові величини

2.1. Випадкові величини, їх види

У багатьох задачах теорії та практики ми маємо справу з числовими величинами, які можуть приймати свої значення в залежності від наслідку випробування. І до того, як випробування проведено, ми не можемо завчасно передбачити, яке саме це буде значення. Наприклад, якщо ми кидаємо гральну кістку (кубик, на кожній грані якого числа від 1 до 6), то ми наперед не можемо сказати, яке саме число очок випаде на верхній грані. Або здійснюється постріл з гвинтівки. Відстань, на яку улетить куля з неї, знову ж таки точно завбачити неможливо – ми можемо тільки вказати деякі межі для цієї відстані. Такі величини називаються випадковими.

Означення. *Випадковою величиною* називається числова величина, яка приймає свої значення в залежності від наслідку випробування, причому в кожному наслідку вона приймає одне й тільки одне значення, наперед не відоме і залежне від випадкових факторів, які завчасно не можуть бути враховані.

Приклади.

1. Нехай ми 10 разів кидаємо монету. Скільки разів випаде «герб»? Завчасно вказати не можна, можна тільки стверджувати, що число появ «герба» буде міститись у межах від 0 до 10. Це число є випадковою величиною.

2. Заміряється відсотковий вміст кальцію у випадково вибраному зразку гірської породи. Зрозуміло, що в різних зразках цей вміст, взагалі кажучи, є різним, і тому завчасно його передбачити неможливо. Цей відсотковий вміст є випадковою величиною.

Вже наведені приклади показують, що випадкові величини можуть бути різних типів. До першого типу можна віднести величини, які можуть приймати тільки окремі, ізольовані одне від одного значення. Наприклад, число очок на верхній грані гральної кістки може приймати лише одне з наступних значень: 1, 2, 3, 4, 5, 6. А число появ «герба» при 10 киданнях монети – лише одне з наступних значень: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Такі випадкові величини називають дискретними.

Означення. *Дискретною випадковою величиною (ДВВ)* називається випадкова величина, яка може приймати тільки окремі, ізольовані одне від одного значення. Число таких значень може бути скінченним, або нескінченним (зліченним), і кожне значення приймається з певною ймовірністю.

Ці значення можна зобразити точками на числовій прямій:

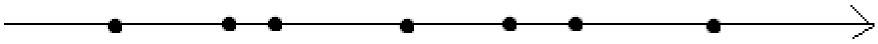


Рис. 12

До іншого типу випадкових величин відносяться так звані неперервні випадкові величини. Це такі величини, можливі значення яких заповнюють собою деякий числовий проміжок, скінченний або нескінченний. До таких величин відносяться, наприклад, відстань, на яку летить куля, відсотковий вміст хімічного елемента у випадково вибраному зразку тощо.

Означення. *Неперервною випадковою величиною (НВВ)* називається випадкова величина, можливі значення якої повністю заповнюють деякий суцільний проміжок, скінченний, або нескінченний.

Рис. 13

Таким чином, якщо множина можливих значень ДВВ скінченна або злічення, то множина значень НВВ континуальна.

Домовимось випадкові величини позначати великими літерами латинського алфавіту: X, Y, …

2.2. Дискретні випадкові величини. Закон розподілу

Тут ми розглянемо більш детально ДВВ. Як можна таку величину охарактеризувати? На першій погляд достатньо перелічити всі її можливі значення. Але в дійсності це не так: дві різні ДВВ можуть мати однакові можливі значення, але ймовірності відповідних значень у цих величин різні. Розглянемо приклад. Нехай монета кидається один раз. ДВВ X – число появ «герба» при цьому киданні. Нехай тепер ми випадковим чином дістаємо карту з колоди в 36 карт. ДВВ Y – число карт пікової масті, що з'явилося. Зрозуміло, що обидві величини X та Y мають однакові можливі значення – 0 («герб» не випав; карта не піка) та 1 («герб» випав; карта – піка). Але ймовірність того, що ДВВ X прийме значення 0, дорівнює $1/2$, а ймовірність того, що ДВВ Y прийме значення 0, дорівнює $3/4$. Відповідно ймовірності значення 1 для цих величин будуть $1/2$ та $1/4$.

Таким чином, ми бачимо, що випадкові величини, які мають однакові можливі значення, можуть мати різні ймовірності цих значень. Тому для того, щоб охарактеризувати ДВВ, необхідно не тільки вказати її можливі значення, а ще ймовірності цих значень.

Означення. *Законом розподілу ДВВ* називається закон відповідності між можливими значеннями ДВВ та ймовірностями цих значень.

Цей закон можна задати таблично, аналітично або графічно. Табличне задання має вигляд таблиці, до якої внесено всі можливі значення ДВВ та ймовірності кожного з цих значень:

X	x_1	x_2	$\ldots$	x_n
P	p_1	p_2	$\ldots$	p_n

Тут $x_1, x_2, \ldots, x_n$ – можливі значення ДВВ. Не обмежуючи загальності вважатимемо, що вони розташовані у порядку зростання: $x_1 < x_2 < \ldots < x_n$. Числа $p_1, p_2, \ldots, p_n$ – відповідні ймовірності цих значень, тобто: $p_j = P(X = x_j)$ $(j = 1, 2, \ldots, n)$. Оскільки в кожному випробуванні випадкова величина може прийняти одне й тільки одне можливе значення, то події $X = x_1$, $X = x_2$, …, $X = x_n$ утворюють повну групу, отже, сума ймовірностей цих подій дорівнює одиниці:

$$p_1 + p_2 + \ldots + p_n = 1. \qquad (*)$$

Рівність (*) називається *умовою нормування*. Якщо множина можливих значень ДВВ нескінченна, то ця умова набуває вигляду:

$$\sum_{j=1}^{\infty} p_j = 1. \qquad (**)$$

Тобто числовий ряд $\sum_{j=1}^{\infty} p_j$ збіжний, та його сума дорівнює 1.

Наприклад, закон розподілу ДВВ, яка дорівнює числу карт пікової масті при випадковому вийманні однієї карти з колоди, має вигляд:

X	0	1
P	$3/4$	$1/4$

Аналітичне задання закону розподілу має вигляд функціональної залежності між значеннями ДВВ та її ймовірностями:

$$p_j = g(x_j) \ \ (j = 1, 2, \ldots, n) \ .$$

Для ДВВ – числа карт пікової масті «серед» однієї карти, що вийнято з колоди, вона може мати вигляд:

$$p_j = -\frac{1}{2}x_j + \frac{3}{4}, \ j = 1, 2,$$

де $x_1 = 0, \ x_2 = 1$.

Графічне задання закону розподілу полягає в побудові так званого *полігону розподілу*. Це ламана лінія на координатній площині, вершини якої знаходяться в точках з координатами $(x_j, p_j) \ \ (j = 1, 2, \ldots, n)$.

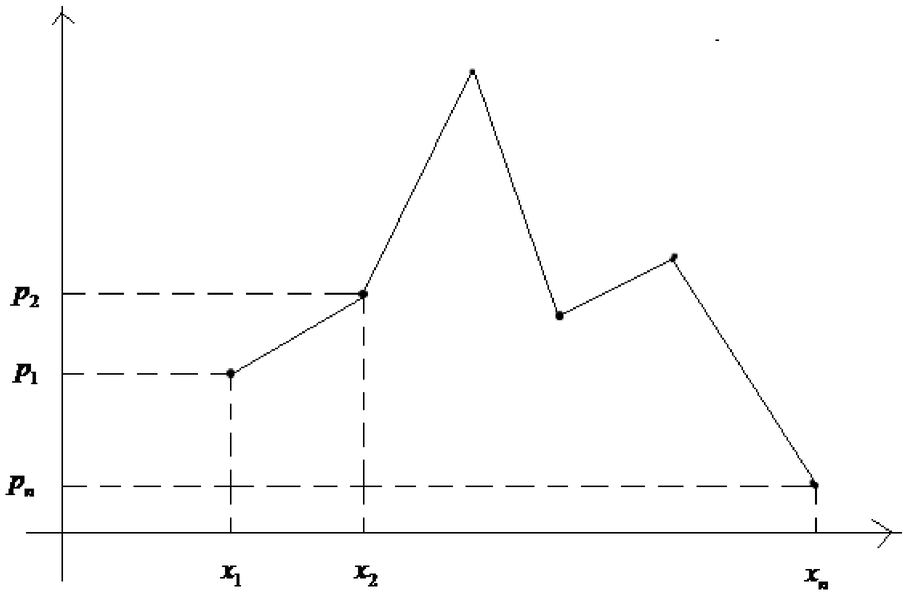


Рис. 14

Нехай, наприклад, закон розподілу ДВВ задано таблицею:

X	2	4	5	6
P	0,3	0,1	0,2	0,4

Тоді полігон розподілу має вигляд:

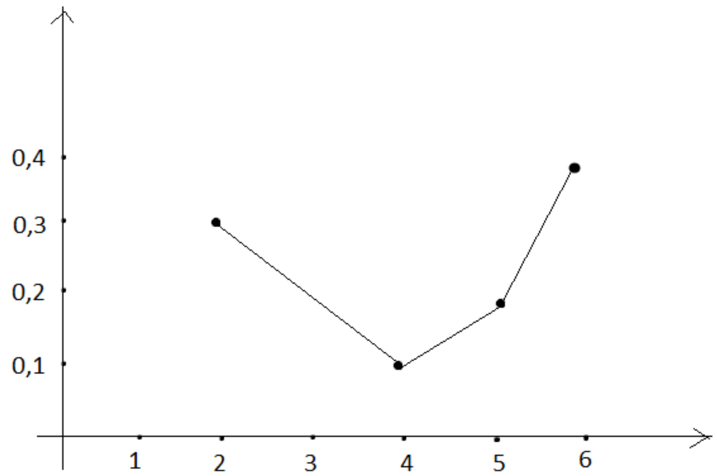


Рис. 15

2.3. Задачі на побудову закону розподілу дискретної випадкової величини

Задача 1. Написати закон розподілу ДВВ – числа очок, яке випадає на верхній грані гральної кістки при одному її киданні.

Дана ДВВ може приймати значення 1, 2, 3, 4, 5, 6, і кожне з цих значень, очевидно, приймається з ймовірністю $1/6$. Отже, ряд розподілу має вигляд:

X	1	2	3	4	5	6
P	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

Задача 2. Два стрілки стріляють по одній цілі по одному разу. Перший стрілок влучає в ціль з ймовірністю 0,9, а другий – з ймовірністю 0,7. Написати закон розподілу ДВВ – числа влучень в ціль.

Очевидно, що кількість влучень в ціль може бути або 0, або 1, або 2. Нехай подія A – влучив перший стрілок, подія B – влучив другий стрілок. Тоді (див. п. 1.10, Задача 5):

$$P(X=0)=P(\overline{A}\overline{B})=P(\overline{A})P(\overline{B})=0{,}1\cdot 0{,}3=0{,}03,$$

$$P(X=1)=P(A\overline{B}+\overline{A}B)=P(A)P(\overline{B})+P(\overline{A})P(B)=0{,}9\cdot 0{,}3+0{,}1\cdot 0{,}7=$$

$$=0{,}27+0{,}07=0{,}34,$$

$$P(X=2)=P(AB)=P(A)P(B)=0{,}9\cdot 0{,}7=0{,}63.$$

Отже закон розподілу:

X	0	1	2
P	0,03	0,34	0,63

Відмітимо, що внаслідок умови нормування останню ймовірність можна було б знайти як різницю між одиницею та сумою попередніх ймовірностей. Але краще її знаходити безпосередньо, а умову нормування використовувати для перевірки: 0,03+0,34+0,63=1.

Задача 3. В партії з 6 деталей є 4 стандартні. Навмання відбирається 3 деталі. Побудувати закон розподілу ДВВ – числа стандартних деталей серед 3-х, що відібрано.

Очевидно, що серед 3-х відібраних може бути або 1, або 2, або 3 стандартні деталі (варіант, коли нема жодної стандартної, неможливий, оскільки тоді вийде, що всі 3 відібрані деталі нестандартні, а нестандартних всього 2). Знайдемо (див. п. 1.5, Задача про вибірку):

$$P(X=k)=\frac{C_4^k C_2^{3-k}}{C_6^3},\ \ k=1,2,3.$$

Маємо:

$$P(X=1)=\frac{C_4^1 C_2^2}{C_6^3}=\frac{4}{20}=\frac{1}{5},$$

$$P(X=2)=\frac{C_4^2 C_2^1}{C_6^3}=\frac{12}{20}=\frac{3}{5},$$

$$P(X=3)=\frac{C_4^3 C_2^0}{C_6^3}=\frac{4}{20}=\frac{1}{5}.$$

Отже закон розподілу:

X	1	2	3
P	$1/5$	$3/5$	$1/5$

2.4. Математичне сподівання дискретної випадкової величини

Не завжди є можливість записати повний закон розподілу випадкової величини – тобто з наведенням всіх можливих значень цієї величини та ймовірностей всіх цих значень. Але для отримання загальної інформації про випадкову величину іноді потреби в такому законі не виникає. Буває достатньо обмежитись деякими числами, які характеризують випадкову величину, так кажучи, сумарно. Такі числа називають *числовими характеристиками* випадкової величини. До них, зокрема, відноситься математичне сподівання.

Означення. *Математичних сподіванням* ДВВ називається сума добутків значень цієї величини на їх ймовірності.

Нехай ДВВ X задано законом розподілу:

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

Тоді математичним сподіванням ДВВ X називається число:

$$M[X] = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n .$$

Нехай ДВВ X приймає зліченну множину значень, тобто має закон розподілу:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

Тоді математичним сподіванням ДВВ X називається сума числового ряду:

$$M[X] = \sum_{j=1}^{\infty} x_j p_j ,$$

за умови, що цей ряд збігається.

Відмітимо, що математичне сподівання ДВВ є величиною невипадковою.

Приклади

1. Знайти математичне сподівання ДВВ, яку задано законом розподілу:

X	-4	6	10
P	0,2	0,3	0,5

Математичне сподівання дорівнює сумі добутків значень ДВВ на їх ймовірності. Маємо:

$$M[X] = -4 \cdot 0,2 + 6 \cdot 0,3 + 10 \cdot 0,5 = 6 .$$

2. Знайти математичне сподівання ДВВ, яку задано законом розподілу:

X	$\frac{1}{1!}$	$\frac{1}{2!}$	$\frac{1}{3!}$	$\cdots$	$\frac{1}{n!}$	$\cdots$
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\cdots$	$\frac{1}{2^n}$	$\cdots$

Дана випадкова величина може приймати зліченну множину значень. Перевіримо спочатку виконання умови нормування. Знайдемо:

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots+\frac{1}{2^n}+\ldots=\frac{1/2}{1-1/2}=1$$

(сума нескінченно спадної геометричної прогресії з першим членом $1/2$ і знаменником також $1/2$). Отже, умову нормування виконано. Знайдемо тепер математичне сподівання:

$$M[X]=1\cdot\frac{1}{2}+\frac{1}{2!}\cdot\left(\frac{1}{2}\right)^2+\frac{1}{3!}\cdot\left(\frac{1}{2}\right)^3+\ldots+\frac{1}{n!}\cdot\left(\frac{1}{2}\right)^n+\ldots=$$

$$=\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{1}{2}\right)^k-1=e^{1/2}-1=\sqrt{e}-1\approx 0,65$$

(на підставі того, що ряд $\sum_{k=0}^{\infty}\frac{x^k}{k!}$ збігається на всій числовій прямій, та його сума дорівнює e^x).

3. Задано перелік можливих значень ДВВ X: $x_1=1$, $x_2=2$, $x_3=3$. Відомо, що $M[X]=2,3$, а $M[X^2]=5,9$. Знайти ймовірності, які відповідають цим можливим значенням.

З умов задачі та з умови нормування маємо:

$$x_1p_1+x_2p_2+x_3p_3=2,3,$$

$$x_1^2p_1+x_2^2p_2+x_3^2p_3=5,9,$$

$$p_1+p_2+p_3=1.$$

Тобто дістали систему 3-х лінійних алгебраїчних рівнянь для невідомих p_1, p_2, p_3. Після підстановки даних задачі ця система набуває вигляду:

$$p_1+2p_1+3p_3=2,3,$$

$$p_1+4p_2+9p_3=5,9,$$

$$p_1+p_2+p_3=1.$$

Розв'язуючи цю систему, отримуємо: $p_1=0,2$, $p_2=0,3$, $p_3=0,5$.

2.5. Основні властивості математичного сподівання

Властивість 1. *Математичне сподівання задовольняє подвійну нерівність:*

$$\min_j x_j \le M[X]\le \max_j x_j.$$

Тобто математичне сподівання ДВВ не може бути менше найменшого її значення і не може бути більше найбільшого її значення.

Доведення. Нехай ДВВ має закон розподілу:

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

Припускаємо, що $x_1 = \min\limits_j x_j$, $x_n = \max\limits_j x_j$. Тоді:

$$M[X] = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n \ge x_1 p_1 + x_1 p_2 + \ldots + x_1 p_n =$$
$$= x_1(p_1 + p_2 + \ldots + p_n) = x_1 \cdot 1 = x_1 = \min_j x_j,$$

$$M[X] = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n \le x_n p_1 + x_n p_2 + \ldots + x_n p_n =$$
$$= x_n(p_1 + p_2 + \ldots + p_n) = x_n \cdot 1 = x_n = \max_j x_j.$$

Властивість 2. *Математичне сподівання сталої величини дорівнює цій самій сталій:*

$M[C] = C$.

Доведення. Сталу C можна розглядати як ДВВ, яка може приймати тільки одне значення C з ймовірністю $p = 1$. А тоді:

$M[C] = C \cdot 1 = C$.

Властивість 3. *Сталий множник можна виносити за знак математичного сподівання:*

$M[CX] = CM[X]$.

Доведення. Нехай ДВВ X має закон розподілу:

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

Тоді ДВВ CX має такий закон розподілу:

CX	Cx_1	Cx_2	$\dots$	Cx_n
P	p_1	p_2	$\dots$	p_n

Знайдемо:

$$M[CX] = Cx_1 \cdot p_1 + Cx_2 \cdot p_2 + \ldots + Cx_n \cdot p_n = C(x_1 p_1 + x_2 p_2 + \ldots x_n p_n) = CM[X].$$

Означення. Дві випадкові величини називаються *незалежними*, якщо закон розподілу кожної з них не залежить від того, які значення прийняла інша величина.

Приклад. Два стрілки стріляють декілька разів по одній цілі. Якщо ДВВ X – число влучень 1-го стрілка, а ДВВ Y – число влучень 2-го стрілка, то величини X та Y – незалежні.

Означення. Якщо випадкові величини X та Y незалежні, то їх *добутком* називається випадкова величина XY, можливі значення якої дорівнюють добуткам кожного можливого значення величини X на кожне можливе значення величини Y; при цьому ймовірності можливих значень величини XY дорівнюють добуткам ймовірностей відповідних значень X та Y.

Властивість 4. *Математичне сподівання добутку двох незалежних ДВВ дорівнює добутку їх математичних сподівань:*

$$M[XY] = M[X]M[Y].$$

Доведення. Для спрощення припустимо, що кожна з величини X та Y приймає тільки два можливі значення з певними ймовірностями, тобто закони їх розподілів мають вигляд:

X	x_1	x_2
P	p_1	p_2

Y	y_1	y_2
P	r_1	r_2

Побудуємо закон розподілу ДВВ XY:

XY	x_1y_1	x_1y_2	x_2y_1	x_2y_2
P	p_1r_1	p_1r_2	p_2r_1	p_2r_2

Знайдемо:

$$M[XY] = x_1y_1p_1r_1 + x_1y_2p_1r_2 + x_2y_1p_2r_1 + x_2y_2p_2r_2 =$$
$$= x_1p_1(y_1r_1 + y_2r_2) + x_2p_2(y_1r_1 + y_2r_2) = (x_1p_1 + x_2p_2)(y_1r_1 + y_2r_2) = M[X]M[Y].$$

Наслідок. *Математичне сподівання добутку декількох взаємно незалежних випадкових величин дорівнює добутку їх математичних сподівань.*

Тобто, якщо $X_1, X_2, \ldots, X_n$ – взаємно незалежні випадкові величини, то

$$M[X_1X_2\cdots X_n] = M[X_1]M[X_2]\cdots M[X_n].$$

Означення. *Сумою* двох випадкових величини X та Y називається випадкова величина $X+Y$, можливі значення якої дорівнюють сумі кожного можливого значення величини X та кожного можливого значення величини Y; якщо величини X та Y незалежні, то ймовірності можливих значень величини $X+Y$ дорівнюють добуткам відповідних значень величин X та Y.

Властивість 5. *Математичне сподівання суми двох випадкових величини дорівнює сумі їх математичних сподівань:*

$$M[X+Y] = M[X] + M[Y].$$

Доведення. Нехай величини X та Y задано законами їх розподілів:

X	x_1	x_2

P	p_1	p_2

Y	y_1	y_2
P	r_1	r_2

(знову для спрощення обмежуємось випадком двох можливих значень кожної з величин). Запишемо закон розподілу величини $X+Y$:

$X+Y$	x_1+y_1	x_1+y_2	x_2+y_1	x_2+y_2
P	p_{11}	p_{12}	p_{21}	p_{22}

Тут $p_{jk}=P(X+Y=x_j+y_j)$. Розглянемо:

$$M[X+Y]=(x_1+y_1)p_{11}+(x_1+y_2)p_{12}+(x_2+y_1)p_{21}+(x_2+y_2)p_{22}=$$
$$=x_1(p_{11}+p_{12})+x_2(p_{21}+p_{22})+y_1(p_{11}+p_{21})+y_2(p_{12}+p_{22}). \qquad (*)$$

Доведемо, що $p_{11}+p_{12}=p_1$. Подія, яка полягає в тому, що величина X прийме значення x_1, еквівалентна події, яка полягає в тому, що величина $X+Y$ прийме значення x_1+y_1 або x_1+y_2. За теоремою про ймовірність суми несумісних подій маємо:

$$P(X=x_1)=P((X+Y=x_1+y_1)+(X+Y=x_1+y_2))=$$
$$=P(X+Y=x_1+y_1)+P(X+Y=x_1+y_2)=p_{11}+p_{12}.$$

Звідси й випливає, що $p_1=p_{11}+p_{12}$. Аналогічно доводимо, що $p_{21}+p_{22}=p_2$, $p_{11}+p_{21}=r_1$, $p_{12}+p_{22}=r_2$. Підставляючи ці рівності в співвідношення (*), дістаємо:

$$M[X+Y]=x_1p_1+x_2p_2+y_1r_1+y_2r_2=M[X]+M[Y].$$

Наслідок 1. *Математичне сподівання суми декількох випадкових величини дорівнює сумі їх математичних сподівань:*

$$M[X_1+X_2+\ldots+X_n]=M[X_1]+M[X_2]+\ldots+M[X_n].$$

Наслідок 2. *Якщо $C_1,C_2,\ldots,C_n$ – сталі величини, то:*

$$M[C_1X_1+C_2X_2+\ldots+C_nX_n]=C_1M[X_1]+C_2M[X_2]+\ldots+C_nM[X_n].$$

Властивість 6. *Математичне сподівання різниці двох випадкових величини дорівнює різниці їх математичних сподівань:*

$$M[X-Y]=M[X]-M[Y].$$

Доведення. Дійсно, на підставі Наслідку 2 з Властивості 5 маємо:

$$M[X-Y]=M[X+(-1)Y]=M[X]+(-1)M[Y]=M[X]-M[Y].$$

Приклад. Знайти математичне сподівання суми очок, які можуть випасти при киданні двох гральних кісток.

Нехай ДВВ X – число очок, що випадає на першій кістці, ДВВ Y – число очок, що випадає на другій кістці. Ці величини мають один й той же закон розподілу (див. п. 2.3, Задача 1). Знайдемо:

$$M[X]=1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=\frac{1}{6}(1+2+3+4+5+6)=\frac{7}{2}.$$

Аналогічно $M[Y]=7/2$. Тоді

$$M[X+Y]=M[X]+M[Y]=\frac{7}{2}+\frac{7}{2}=7.$$

2.6. Ймовірнісний зміст математичного сподівання

Нехай проведено n випробувань, у кожному з яких випадкова величина X приймає яке-небудь значення. Припустимо, що m_1 разів величина X прийняла значення x_1; m_2 разів прийняла значення x_2; … ; m_k разів прийняла значення x_k; $m_1+m_2+\ldots+m_k=n$. Знайдемо середнє арифметичне всіх значень, що прийнято величиною X:

$$\overline{x}=\frac{x_1m_1+x_2m_2+\ldots+x_km_k}{n}=x_1\left(\frac{m_1}{n}\right)+x_2\left(\frac{m_2}{n}\right)+\ldots+x_k\left(\frac{m_k}{n}\right).$$

Частка m_j/n є відносною частотою w_j значення x_j $(j=1,2,...,k)$. З урахуванням цього величина $\overline{x}$ запишеться так:

$$\overline{x}=x_1w_1+x_2w_2+\ldots+x_kw_k.$$

Пригадаємо, що при збільшенні n відносна частота події, як правило, наближається до ймовірності цієї події. Тому наближено можемо покласти:

$$w_1\approx p_1, w_2\approx p_2,..., w_k\approx p_k,$$

де $p_j=P(X=x_j)$ $(j=1,2,...,k)$. Тоді

$$\overline{x}\approx x_1p_1+x_2p_2+\ldots+x_kp_k=M[X].$$

Таким чином, ймовірнісний зміст математичного сподівання такий: *математичне сподівання наближено дорівнює (тим точніше, чим більше число випробувань) середньому арифметичному значень випадкової величини, що нею приймаються.*

Приклад. Два гравці A і B грають за наступними правилами. Вони кидають гральну кістку. Якщо випадає «6», то гравець A платить гравцю B 6 у. о. Але якщо «6» не випадає, то гравець B платить гравцю A 1 у. о. Для якого з гравців ця гра більш вигідна?

Побудуємо закони розподілів випадкових величин X (виграш гравця A) та Y (виграш гравця B). Гравець A може за один кон або придбати 1 у. о., або втратити 6 у. о. (тобто «придбати» -6 у. о.). Ймовірність того, що він придбає 1 у. о., дорівнює ймовірності випадіння будь якого числа очок, крім «6», тобто $5/6$. А ймовірність втрати 6 у. о. дорівнює ймовірності випадіння «6», тобто $1/6$. Отже закон розподілу виграшу для гравця A такий:

X	-6	1
P	$1/6$	$5/6$

Аналогічно для гравця B матимемо:

Y	-1	6
P	$5/6$	$1/6$

Тоді

$$M[X] = -6 \cdot \frac{1}{6} + 1 \cdot \frac{5}{6} = -\frac{1}{6}, \quad M[Y] = -1 \cdot \frac{5}{6} + 6 \cdot \frac{1}{6} = \frac{1}{6} .$$

Таким чином, ми бачимо, що така гра вигідна для гравця B. Адже математичне сподівання його виграшу додатне, а для гравця A – від'ємне. Це означає, що якщо ця гра продовжуватиметься досить довго, то гравець B буде, скоріше всього, у виграші, а гравець A – у програші.

2.7. Дисперсія дискретної випадкової величини

Однієї такої характеристики, як математичне сподівання, буває недостатньо для інформативного опису випадкової величини. Розглянемо наступний приклад. Нехай задано дві ДВВ X та Y своїми законами розподілів:

X	-0,01	0,01
P	0,5	0,5

Y	-100	100
P	0,5	0,5

Знайдемо:

$$M[X] = -0{,}01 \cdot 0{,}5 + 0{,}01 \cdot 0{,}5 = 0, \quad M[Y] = -100 \cdot 0{,}5 + 100 \cdot 0{,}5 = 0 .$$

Таким чином, ми бачимо, що ці величини мають однакові математичні сподівання. Але розсіяння значень цих величин навколо цього математичного сподівання суттєво розрізняються. Якщо величина X має значення досить близькі до математичного сподівання, то значення величини Y розташовані від нього значно дальше. Тому для більш інформативного опису випадкової величини треба вводити характеристику, яка б враховувала це розсіяння. Наприклад, в артилерії необхідно знати, наскільки купно влучать снаряди поблизу цілі. З цією метою вводиться така числова характеристика випадкової величини, як дисперсія.

Означення. *Дисперсією* випадкової величини називається математичне сподівання квадрату відхилення випадкової величини від її математичного сподівання:

$$D[X] = M\left[(X - M[X])^2\right].$$

Приклад. Знайти дисперсію ДВВ X, яку задано законом розподілу:

X	-5	2	3	4
P	0,4	0,3	0,1	0,2

Маємо:

$$M[X] = -5 \cdot 0{,}4 + 2 \cdot 0{,}3 + 3 \cdot 0{,}1 + 4 \cdot 0{,}2 = -0{,}3.$$

Побудуємо закон розподілу величини $(X - M[X])^2$:

$(X - M[X])^2$	22,09	5,29	10,89	18,49
P	0,4	0,3	0,1	0,2

(для отримання значень цієї величини від кожного значення величини X відняли математичне сподівання і результат підвели до квадрату; ймовірності залишаються тими ж самими). Звідси маємо:

$$D[X] = M\left[(X - M[X])^2\right] =$$

$$= 22{,}09 \cdot 0{,}4 + 5{,}29 \cdot 0{,}3 + 10{,}89 \cdot 0{,}1 + 18{,}49 \cdot 0{,}2 = 15{,}21.$$

На практиці дисперсію часто (але не завжди) зручніше знаходити за наступною формулою:

$$D[X] = M[X^2] - M^2[X]. \qquad (*)$$

Тобто дисперсія дорівнює математичному сподіванню квадрату випадкової величини без квадрату її математичного сподівання. Дійсно, розглянемо:

$$D[X] = M\left[(X - M[X])^2\right] = M\left[X^2 - 2M[X]X + M^2[X]\right] =$$

$$= M[X^2] - 2M[X]M[X] + M[M^2[X]] =$$

$$= M[X^2] - 2M^2[X] + M^2[X] = M[X^2] - M^2[X].$$

Тут ми скористалися наступними властивостями математичного сподівання: $M[C] = C$, $M[CX] = CM[X]$, $M[X \pm Y] = M[X] \pm M[Y]$.

Приклад. Знайти дисперсію ДВВ, яку задано законом розподілу:

X	-1	0	3	5	6
P	0,1	0,1	0,3	0,2	0,3

Скористаємось формулою (*). Маємо:

$M[X] = -1\cdot 0{,}1 + 0\cdot 0{,}1 + 3\cdot 0{,}3 + 5\cdot 0{,}2 + 6\cdot 0{,}3 = 3{,}6$;

$M[X^2] = 1\cdot 0{,}1 + 0\cdot 0{,}1 + 9\cdot 0{,}3 + 25\cdot 0{,}2 + 36\cdot 0{,}3 = 18{,}6$;

$D[X] = M[X^2] - M^2[X] = 18{,}6 - 3{,}6^2 = 5{,}64$.

Розглянемо основні властивості дисперсії.

Властивість 1. *Дисперсія завжди є числом невід'ємним:*

$D[X] \ge 0$.

Дійсно, величина $(X - M[X])^2$ невід'ємна, отже математичне сподівання цієї величини також невід'ємне.

Властивість 2. *Дисперсія сталої величини дорівнює нулю.*

Дійсно, оскільки $M[C] = C$, то

$$D[C] = M\left[\left(C - M[C]\right)^2\right] = M\left[(C - C)^2\right] = M[0] = 0.$$

Стала величина зберігає одне й те ж значення, тому розсіяння не має.

Властивість 3. *Сталий множник можна виносити за знак дисперсії, підводячи при цьому його до квадрату:*

$$D[CX] = C^2 D[X].$$

Дійсно, розглянемо:

$$D[CX] = M\left[(CX - M[CX])^2\right] = M\left[(CX - CM[X])^2\right] =$$

$$= M\left[C^2(X - M[X])^2\right] = C^2 M\left[(X - M[X])^2\right] = C^2 D[X].$$

Властивість 4. *Дисперсія суми двох незалежних випадкових величин дорівнює сумі їх дисперсій:*

$$D[X+Y] = D[X] + D[Y].$$

Доведення. За формулою (*) маємо:

$$D[X+Y] = M\left[(X+Y)^2\right] - M^2[X+Y] = M\left[X^2 + 2XY + Y^2\right] - \left(M[X] + M[Y]\right)^2 =$$

$$= M[X^2] + 2M[X]M[Y] + M[Y^2] - M^2[X] - 2M[X]M[Y] - M^2[Y] =$$

$$= M[X^2] - M[X^2] + M[Y^2] - M^2[Y] = D[X] + D[Y].$$

Наслідок 1. *Дисперсія суми декількох взаємно незалежних випадкових величин дорівнює сумі їх дисперсій:*

$$D[X_1 + X_2 + \ldots + X_n] = D[X_1] + D[X_2] + \ldots + D[X_n].$$

Наслідок 2. *Додавання до випадкової величини сталої не змінює дисперсії випадкової величини:*

$D[X+C]=D[X]$.

Дійсно:

$D[X+C]=D[X]+D[C]=D[X]$.

Властивість 5. *Дисперсія різниці двох незалежних випадкових величин дорівнює сумі їх дисперсій:*

$D[X-Y]=D[X]+D[Y]$.

Дійсно:

$D[X-Y]=D[X+(-1)Y]=D[X]+D[(-1)^2Y]=D[X]+(-1)^2D[Y]=D[X]+D[Y]$.

Звернемо увагу на відмінність цієї властивості від відповідної властивості математичного сподівання: $M[X-Y]=M[X]-M[Y]$.

2.8. Середньоквадратичне відхилення дискретної випадкової величини

Використання дисперсії для врахування розсіяння значень випадкової величини навколо математичного сподівання має ту незручність, що змінює розмірність випадкової величини, адже визначається через її квадрат. Наприклад, якщо випадкова величина вимірюється в сантиметрах, то її дисперсія буде вимірюватись в квадратних сантиметрах. З метою уникнення такої незручності користуються іншою характеристикою розсіяння.

Означення. *Середньоквадратичним відхиленням* $\sigma[X]$ випадкової величини X називається арифметичний квадратний корінь з її дисперсії:

$\sigma[X]=\sqrt{D[X]}$.

Приклад. Знайти середньоквадратичне відхилення ДВВ X, яку задано законом розподілу:

X	4,3	5,1	10,6
P	0,2	0,3	0,5

Маємо:

$M[X]=4,3\cdot 0,2+5,1\cdot 0,3+10,6\cdot 0,5=7,69$;

$M[X^2]=4,3^2\cdot 0,2+5,1^2\cdot 0,3+10,6^2\cdot 0,5=67,681$;

$D[X]=M[X^2]-M^2[X]=67,681-7,69^2=8,5449$;

$\sigma[X]=\sqrt{D[X]}=\sqrt{8,5449}\approx 2,923$.

Теорема. *Середньоквадратичне відхилення суми скінченного числа взаємно незалежних випадкових величин дорівнює арифметичному квадратному кореню з суми квадратів середньоквадратичних відхилень цих величин:*

$$\sigma[X_1+X_2+...+X_n]=\sqrt{\sigma^2[X_1]+\sigma^2[X_2]+...+\sigma^2[X_n]}\,.$$

Доведення. Позначимо: $X=X_1+X_2+...+X_n$. Тоді (див. п. 2.7, наслідок 1 з Властивості 4):

$$D[X]=D[X_1+X_2+...+X_n]=D[X_1]+D[X_2]+\ldots+D[X_n].$$

Звідси:

$$\sigma[X]=\sqrt{D[X]}=\sqrt{D[X_1]+D[X_2]+\ldots+D[X_n]}=\sqrt{\sigma^2[X_1]+\sigma^2[X_2]+...+\sigma^2[X_n]}\,.$$

2.9. Числові характеристики середнього арифметичного однаково розподілених випадкових величин

Нехай $X_1,X_2,...,X_n$ – взаємно незалежні випадкові величини, які мають один й той же розподіл, отже однакові числові характеристики (математичне сподівання та дисперсію):

$$M[X_k]=a,\ D[X_k]=D,\ \sigma[X_k]=\sigma=\sqrt{D}\quad (k=1,2,...,n)\,.$$

Позначимо:

$$\overline{X}_n=\frac{X_1+X_2+...+X_n}{n}=\frac{1}{n}(X_1+X_2+...+X_n).$$

Тобто величина $\overline{X}_n$ є середнє арифметичне випадкових величин $X_1,X_2,...,X_n$. Нашою метою буде встановлення зв'язку між числовими характеристиками величини $\overline{X}_n$ та числовими характеристиками величин $X_1,X_2,...,X_n$.

Почнемо з математичного сподівання. Розглянемо:

$$M\left[\overline{X}_n\right]=M\left[\frac{1}{n}(X_1+X_2+...+X_n)\right]=\frac{1}{n}M[X_1+X_2+...+X_n]=$$

$$=\frac{1}{n}(M[X_1]+M[X_2]+\ldots+M[X_n])=\frac{1}{n}\left(\underbrace{a+a+\ldots+a}_{n\ \text{разів}}\right)=\frac{1}{n}na=a\,.$$

Математичне сподівання середнього арифметичного однаково розподілених взаємно незалежних випадкових величин дорівнює математичному сподіванню кожної з цих величин.

Розглянемо тепер дисперсію:

$$D\left[\bar{X}_n\right]=D\left[\frac{1}{n}(X_1+X_2+\ldots+X_n)\right]=\frac{1}{n^2}D[X_1+X_2+\ldots+X_n]=$$

$$=\frac{1}{n^2}(D[X_1]+D[X_2]+\ldots+D[X_n])=\frac{1}{n^2}\left(\underbrace{D+D+\ldots+D}_{n\ разів}\right)=\frac{1}{n^2}nD=\frac{D}{n}.$$

Дисперсія середнього арифметичного n однаково розподілених взаємно незалежних випадкових величин в n разів менше дисперсії D кожної з цих величин.

Розглянемо середньоквадратичне відхилення:

$$\sigma\left[\bar{X}_n\right]=\sqrt{D\left[\bar{X}_n\right]}=\sqrt{\frac{D}{n}}=\frac{\sqrt{D}}{\sqrt{n}}=\frac{\sigma}{\sqrt{n}}.$$

Середньоквадратичне відхилення середнього арифметичного n однаково розподілених взаємно незалежних випадкових величин в $\sqrt{n}$ разів менше середньоквадратичного відхилення σ кожної з цих величин.

Об'єднуючи ці результати, можна зробити висновок, що середнє арифметичне досить великого числа однаково розподілених взаємно незалежних випадкових величин має значно менше розсіяння, ніж кожна з цих величин окремо. Цей результат має важливе прикладне значення. Нехай ми проводимо вимірювання деякої фізичної величини. Результат кожного вимірювання залежить від багатьох випадкових факторів, отже є величиною випадковою. Для знаходження значення шуканої фізичної величини, як правило, проводять не одне, а декілька вимірювань, а потім знаходять середнє арифметичне результатів всіх вимірювань. Чому так робиться? А ось як раз на підставі властивостей числових характеристик величини $\bar{X}_n$. Результати вимірювань, як випадкові величини, мають однаковий розподіл (якщо вимірювання проводяться одними й тими ж приладами, за однією і тою ж методикою), отже, однакові числові характеристики. Вони взаємно незалежні (результат кожного вимірювання не залежить від інших вимірювань). Тоді середнє арифметичне результатів вимірювань матиме менше розсіяння, ніж кожне вимірювання окремо. Це означає, що середнє арифметичне має меншу похибку, отже, дає більш точне наближення до шуканого результату, ніж результат одного окремого вимірювання. І це наближення буде тим точніше, чим більше вимірювань проведено.

2.10. Початкові та центральні теоретичні моменти

Розглянемо ДВВ X, яку задано наступним законом розподілу:

X	1	2	5
P	0,6	0,2	0,2

Її математичне сподівання:

$$M[X]=1\cdot 0,6+2\cdot 0,2+5\cdot 0,2=2 \ .$$

А тепер розглянемо іншу ДВВ Y, яка, порівняно з величиною X, має ще одне можливе значення, яке значно більше значень величини X:

Y	1	2	5	100
P	0,6	0,2	0,19	0,01

Ймовірність цього значення 100 дуже мала. Тобто при випробуваннях таке значення приймається випадковою величиною дуже рідко. Знайдемо:

$$M[Y]=1\cdot 0,6+2\cdot 0,2+5\cdot 0,19+100\cdot 0,01=2,95 .$$

Ми бачимо, що, незважаючи на появу у величини Y нового значення 100, яке значно більше решти значень, математичне сподівання величини Y, порівняно з математичним сподіванням величини X, змінилося далеко не так значно. Тобто це велике значення 100 у математичному сподівання величини Y практично не урахувалося – число $M[Y]=2,95$ за своїм порядком таке ж саме, як і число $M[X]=2$. Це пов'язано з тим, що ймовірність значення 100 мала порівняно з ймовірностями інших значень.

Розглянемо тепер

$$M\left[Y^2\right]=1\cdot 0,6+4\cdot 0,2+25\cdot 0,19+10000\cdot 0,01=106,15 .$$

А ось це число вже значно більше, ніж $M[Y]$, і за своїм порядком воно таке ж саме, як це рідке значення 100. Тобто число $M\left[Y^2\right]$ дозволяє краще урахувати вплив на математичне сподівання того можливого значення, яке велике, але має малу ймовірність. Тому доцільно розглядати математичне сподівання натурального степеня випадкової величини.

Означення. *Початковим моментом* ν_k *порядку* k випадкової величини X називається математичне сподівання величини X^k:

$$\nu_k=M[X^k].$$

Зокрема:

$$\nu_1=M[X],\ \ \nu_2=M[X^2].$$

З використанням початкових моментів формулу (*) п. 2.7 для обчислення дисперсії можна записати так:

$$D[X] = \nu_2 - \nu_1^2 .$$

Означення. *Центральним моментом* μ_k *порядку* k випадкової величини X називається математичне сподівання величини $(X - M[X])^k$:

$$\mu_k = M\left[(X - M[X])^k \right] .$$

Зокрема:

$$\mu_1 = M[X - M[X]] = M[X] - M[M[X]] = M[X] - M[X] = 0 ,$$

$$\mu_2 = M\left[(X - M[X])^2 \right] = D[X].$$

Таким чином, маємо:

$$\mu_2 = \nu_2 - \nu_1^2 .$$

Нескладно довести також справедливість наступних формул:

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3 ,$$

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_4\nu_1^2 - 3\nu_1^4 .$$

Моменти вищих порядків використовуються рідко.

Приклад. ДВВ X задано законом розподілу:

X	3	5
P	0,2	0,8

Знайти центральні моменти 1-го, 2-го, 3-го та 4-го порядків.

Знайдемо спочатку початкові моменти 1-го, 2-го, 3-го та 4-го порядків:

$$\nu_1 = M[X] = 3 \cdot 0,2 + 5 \cdot 0,8 = 4,6 ;$$

$$\nu_2 = M[X^2] = 9 \cdot 0,2 + 25 \cdot 0,8 = 21,8 ;$$

$$\nu_3 = M[X^3] = 27 \cdot 0,2 + 125 \cdot 0,8 = 105,4 ;$$

$$\nu_4 = M[X^4] = 81 \cdot 0,2 + 625 \cdot 0,8 = 516,2 .$$

Тепер знайдемо центральні моменти:

$$\mu_1 = 0 ;$$

$$\mu_2 = \nu_2 - \nu_1^2 = 21,8 - 4,6^2 = 0,64 ;$$

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3 = 105,4 - 3 \cdot 21,8 \cdot 4,6 + 2 \cdot 4,6^2 = -0,768 ;$$

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_4\nu_1^2 - 3\nu_1^4 =$$

$$= 516,2 - 4 \cdot 105,4 \cdot 4,6 + 6 \cdot 516,2 \cdot 4,6^2 - 3 \cdot 4,6^4 \approx 62,77 .$$

2.11 Біноміальний розподіл

У цьому та двох наступних параграфах ми познайомимось з деякими законами розподілів дискретних випадкових величин, які зустрічаються в прикладних задачах. Тут ми розглянемо так званий біноміальний розподіл.

Означення. ДВВ X називається розподіленою за *біноміальним законом* з параметром p $(0 < p < 1)$, якщо вона приймає значення $0,1,2,\ldots,n$ (n – натуральне число), причому $P(X = k) = C_n^k p^k (1-p)^{n-k} \quad (k = 0,1,2,\ldots,n)$.

Тобто закон розподілу величини X має вигляд:

X	0	1	2	$\cdots$	k	$\cdots$	n
P	p_0	p_1	p_2	$\cdots$	p_k	$\cdots$	p_n

$$p_k = C_n^k p^k (1-p)^{n-k} \quad (k = 0,1,2,\ldots,n).$$

Легко помітити, що ДВВ X – число появ у n випробуваннях події A за умови, що ймовірність появи події A у одному випробуванні дорівнює p. Та ймовірність того, що подія A з'явиться k разів, тобто $P(X = k)$, обчислюється за формулою Бернуллі (див. п. 1.13). Біноміально розподілені величини досить поширені на практиці. Зокрема, вони особливо важливі в ігрових задачах.

Знайдемо числові характеристики величини X. Перевіримо спочатку виконання умови нормування. Розглянемо:

$$\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1.$$

Тут ми використали формулу бінома Ньютона:

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

(звідси й назва розподілу – біноміальний).

Знайдемо математичне сподівання величини X. Тут можна використати два шляхи. Перший полягає у розгляді формального виразу для математичного сподівання величини X:

$$M[X] = \sum_{k=0}^{n} k p_k = \sum_{k=0}^{n} k C_n^k p^k (1-p)^{n-k}$$

та обчислення цієї суми. Але цей спосіб складний. Краще використати інший, який ґрунтується саме на тому, що величина X – число появ події A в n випробуваннях.

Розглянемо випадкову величину I_A, яку задано наступним законом розподілу:

I_A	0	1
P	q	p

Тут $q = 1 - p$. Така величина називається *індикатором події* A. Вона дорівнює числу появ події A в одному випробуванні, якщо $P(A) = p$. Знайдемо:

$$M[I_A] = M[I_A^2] = 0 \cdot q + 1 \cdot p = p,$$

$$D[I_A] = M[I_A^2] - M^2[I_A] = p - p^2 = p(1-p) = pq.$$

Повернемось до величини X. Позначимо як X_k – число появ події A у k-му випробуванні. Очевидно, що всі величини $X_1, X_2, ..., X_n$ взаємно незалежні та мають один й той же закон розподілу, який збігається з законом розподілу величини I_A. Тому $M[X_k] = p$, $D[X_k] = pq$ $(k = 1, 2, ..., n)$. Крім того, очевидно, що

$$X = X_1 + X_2 + \ldots + X_n.$$

Тому:

$$M[X] = M[X_1 + X_2 + \ldots + X_n] = M[X_1] + M[X_2] + \ldots + M[X_n] = \underbrace{p + p + \ldots p}_{n \text{ разів}} = np,$$

$$D[X] = D[X_1 + X_2 + \ldots + X_n] = D[X_1] + D[X_2] + \ldots + D[X_n] =$$
$$= \underbrace{pq + pq + \ldots pq}_{n \text{ разів}} = npq.$$

<u>Приклад.</u> Проводиться 10 кидань гральної кістки. Знайти математичне сподівання та дисперсію числа появ «шестірки».

Маємо: $p = 1/6$, $q = 5/6$;

$$M[X] = np = 10 \cdot \frac{1}{6} = \frac{10}{6} = \frac{5}{3}, \quad D[X] = npq = 10 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{18}.$$

2.12. Розподіл Пуассона

Означення. ДВВ X називається розподіленою *за законом Пуассона* з параметром $\lambda > 0$, якщо вона може приймати значення $0, 1, 2, \ldots, k, \ldots$, причому

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

Тобто закон розподілу має вигляд:

X	0	1	2	$\ldots$	k	$\ldots$

P	p_0	p_1	p_2	$\cdots$	p_k	$\cdots$

$$p_k = \frac{\lambda^k}{k!}e^{-\lambda} \quad (k = 0,1,2,...).$$

Ця величина може приймати зліченну множину значень. Розподіл Пуассона іноді називають розподілом рідких подій. Прикладами випадкових величин, розподілених за цим законом можуть бути: число нещасних випадків, число дефектів у виробничих процесах тощо. Цей розподіл використовується у задачах статистичного контролю якості, в теорії надійності, теорії масового обслуговування та інших галузях.

Якщо у схемі незалежних випробувань число випробувань n досить велике, а ймовірність p появи події в одному випробуванні прямує до нуля, то біноміальний розподіл є близьким до розподілу Пуассона, параметр якого $\lambda = np$ (див. п. 1.14).

Перевіримо виконання умови нормування. Розглянемо:

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}\cdot e^{\lambda} = 1.$$

Знайдемо тепер математичне сподівання розподілу Пуассона:

$$M[X] = \sum_{k=0}^{\infty} kp_k = \sum_{k=1}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{s=0}^{\infty}\frac{\lambda^s}{s!} = \lambda e^{-\lambda}\cdot e^{\lambda} = \lambda.$$

Отже математичне сподівання розподілу Пуассона дорівнює параметру λ.

Знайдемо тепер дисперсію. Для цього спочатку знайдемо математичне сподівання квадрату випадкової величини:

$$M[X^2] = \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=1}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{k\lambda^{k-1}}{(k-1)!} =$$

$$= \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{(k-1+1)\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\left(\sum_{k=1}^{\infty}\frac{(k-1)\lambda^{k-1}}{(k-1)!} + \sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) =$$

$$= \lambda e^{-\lambda}\left(\lambda\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) = \lambda e^{-\lambda}(\lambda e^{-\lambda} + e^{\lambda}) = \lambda^2 + \lambda.$$

Звідси маємо:

$$D[X] = M[X^2] - M^2[X] = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Таким чином, дисперсія розподілу Пуассона співпадає з математичним сподіванням і також дорівнює параметру λ. Очевидно тоді, що середньоквадратичне відхилення

$$\sigma[X] = \sqrt{\lambda}.$$

Приклад. Прилад складається з великого числа незалежно працюючих елементів з однаковою дуже малою ймовірністю відмови кожного елементу за час T. Знайти середнє число елементів, що відмовили за час T, якщо ймовірність того, що за цей час відмовить хоча б один елемент, дорівнює 0,98.

Оскільки число елементів велике, елементи працюють незалежно один від одного, а ймовірність відмови кожного елемента мала, то число відмовлень розподілено за законом Пуассона. І треба знайти параметр λ – середнє число відмовлень.

Ймовірність того, що відмовить хоча б один елемент, дорівнює (див. п. 1.14):

$$1-\frac{\lambda^0}{0!}e^{-\lambda}=1-e^{-\lambda},$$

отже

$$1-e^{-\lambda}=0{,}98,$$

звідки

$$e^{-\lambda}=1-0{,}98=0{,}02.$$

Тоді

$$\lambda=-\ln 0{,}02\approx 3{,}91.$$

Отже за час T відмовить у середньому 4 елементи.

2.13. Геометричний розподіл

Означення. ДВВ X називається розподіленою за *геометричним законом* з параметром p $(0<p<1)$, якщо вона може приймати значення $1,2,...,k,...$, причому $P(X=k)=q^{k-1}p$, де $q=1-p$.

Тобто закон розподілу має вигляд:

X	1	2	3	$\cdots$	k	$\cdots$
P	p	qp	q^2p	$\cdots$	$q^{k-1}p$	$\cdots$

Припустимо, що ймовірність появи події A в одному випробуванні дорівнює p. Скільки треба провести випробувань до першої появи події A? Число цих випробувань й є випадкова величина X, яка розподілена за геометричним законом. Очевидно, що вона може приймати зліченну множину значень: теоретично можна нескінченно продовжувати випробування, доки подія A з’явиться.

Перевіримо виконання умови нормування. Знайдемо:

$$\sum_{k=1}^{\infty} P(X=k) = p + qp + q^2 p + \ldots + q^{k-1} p + \ldots = \frac{p}{1-q} = \frac{p}{p} = 1 .$$

Тут ми скористалися формулою суми нескінченно спадної геометричної прогресії (звідси й назва розподілу) з першим членом p і знаменником $q = 1 - p$.

Знайдемо тепер математичне сподівання:

$$M[X] = \sum_{k=1}^{\infty} k p_k = \sum_{k=1}^{\infty} k q^{k-1} p = p \sum_{k=1}^{\infty} k q^{k-1} = p \sum_{k=1}^{\infty} \frac{d}{dq}\left(q^k\right) =$$

$$= p \frac{d}{dq}\left(\sum_{k=1}^{\infty} q^k\right) = p \frac{d}{dq}\left(\frac{q}{1-q}\right) = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} .$$

Тут ми скористалися тим, що степеневий ряд на інтервалі його збіжності можна почленно диференціювати.

Таким чином, середнє число випробувань до першої появи події обернено пропорційно ймовірності цієї події.

Знайдемо дисперсію:

$$D[X] = \sum_{k=1}^{\infty} k^2 p_k - \left(\frac{1}{p}\right)^2 = \sum_{k=1}^{\infty} k^2 q^{k-1} p - \frac{1}{p^2} = p \sum_{k=1}^{\infty} k^2 q^{k-1} - \frac{1}{p^2} .$$

Розглянемо спочатку:

$$\sum_{k=1}^{\infty} k^2 q^{k-1} = 1 + \sum_{k=2}^{\infty} k^2 q^{k-1} = 1 + q \sum_{k=2}^{\infty} k^2 q^{k-2} = 1 + q\left[\sum_{k=2}^{\infty}\left(k^2 - k\right) q^{k-2} + \sum_{k=2}^{\infty} k q^{k-2}\right] =$$

$$= 1 + q\left[\sum_{k=2}^{\infty} k(k-1) q^{k-2} + \frac{1}{q}\left(\sum_{k=1}^{\infty} k q^{k-1} - 1\right)\right] =$$

$$= 1 + q\left[\frac{d^2}{dq^2}\left(\sum_{k=0}^{\infty} q^k\right) + \frac{1}{q}\left(\frac{d}{dq}\left(\sum_{k=0}^{\infty} q^k\right) - 1\right)\right] = 1 + q\left[\frac{2}{(1-q)^3} + \frac{1}{q}\left(\frac{1}{(1-q)^2} - 1\right)\right] =$$

$$= 1 + \frac{2q}{(1-q)^3} + \frac{1}{(1-q)^2} - 1 = \frac{1+q}{(1-q)^3} = \frac{1+q}{p^3} .$$

Тепер з виразу для дисперсії маємо:

$$D[X] = p \frac{1+q}{p^3} - \frac{1}{p^2} = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2} .$$

Тоді середньоквадратичне відхилення:

$$\sigma[X] = \sqrt{D[X]} = \sqrt{\frac{q}{p^2}} = \frac{\sqrt{q}}{p} .$$

Приклад. Скільки в середньому треба провести пострілів до цілі, якщо кожне влучання відбувається з ймовірністю 0,02?

Скористаємось геометричним розподілом. Середнє число випробувань (математичне сподівання) для нього дорівнює $1/p$, отже

$$\overline{X} \approx \frac{1}{0,02} = 50 .$$

Тобто в середньому треба 50 пострілів.

2.14. Неперервні випадкові величини. Функція розподілу

Перейдемо тепер до розгляду неперервних випадкових величин (НВВ). Знову виникає питання – як таку величину можна охарактеризувати? Зрозуміло, що задати перелік всіх її можливих значень, як ми це робили для дискретних випадкових величин, тепер неможливо. Адже неможливо перелічити всі дійсні числа з суцільного проміжку. Тому виникає необхідність надання загального способу опису випадкових величин будь-яких типів. Таким способом є введення так званої функції розподілу випадкової величини.

Нехай x – дійсне число, X – випадкова величина. При проведенні випробувань величина X може з певною ймовірністю прийняти значення, яке менше x. Очевидно, що, якщо x змінюється, то, взагалі кажучи, буде змінюватися й ця ймовірність, тобто ця ймовірність є функцією x. Ця функція називається функцією розподілу випадкової величини X.

Означення. *Функцією розподілу* випадкової величини X називається функція $F(x)$, яка дорівнює ймовірності того, що внаслідок випробування випадкова величина X прийме значення, менше, ніж x:

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

З геометричної точки зору функція $F(x)$ є ймовірність того, що випадкова величина X внаслідок випробування прийме значення, яке на числовій прямій зображується точкою, що лежить зліва від точки x, тобто потрапить до інтервалу $(-\infty, x)$ (рис. 16).

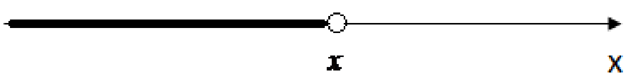


Рис. 16

<u>Приклад</u>. Нехай точка M випадковим чином потрапляє в круг радіусу R. Нехай X – випадкова величина, яка дорівнює відстані від точки M до центра круга (рис. 17). Побудувати функцію розподілу величини X.

Нехай x – довільне дійсне число. Для кожного значення x нам треба знайти $P(X < x)$. Розглянемо три можливі випадки.

1). $x \le 0$. Тоді, оскільки X – це за умовою задачі відстань, а відстань не може бути від'ємною, подія $X < x$ неможлива, отже $F(x) = P(X < x) = 0$.

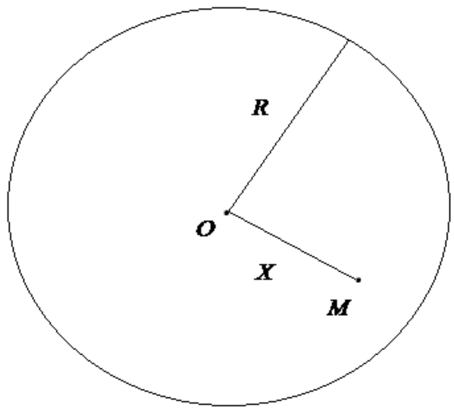


Рис. 17

2). $0 < x \le R$. Тоді $P(X < x)$ є ймовірністю того, що точка M, потрапляючи в круг радіусу R, потрапить при цьому в концентричний з ним менший (принаймні, не більший) круг радіусу $x \le R$. Згідно з формулою геометричної ймовірності (див. п. 1.7):

$$F(x) = P(X < x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}.$$

3). $x > R$. Тоді $P(X < x)$ є ймовірністю того, що точка M, потрапляючи в круг радіусу R, потрапить в концентричний з ним більший круг радіусу $x > R$. А це трапиться обов'язково, тобто подія $X < x$ цього разу достовірна, тому $F(x) = P(X < x) = 1$.

Таким чином, маємо:

$$F(x) = \begin{cases} 0, \ x \le 0, \\ x^2/R^2, \ 0 < x \le R, \\ 1, \ x > R. \end{cases}$$

Графік функції $F(x)$ наведено на рис. 18.

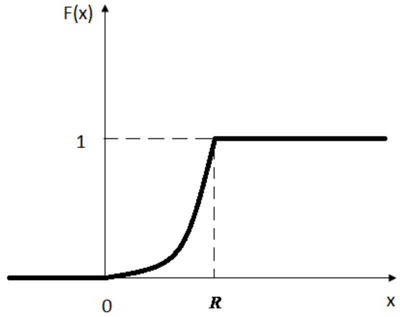


Рис. 18

2.15. Властивості функції розподілу

Властивість 1. *Значення функції розподілу належать відрізку* $[0,1]$:
$0 \le F(x) \le 1$.

Дійсно, за означенням функція розподілу є ймовірність, а ймовірність завжди є число з відрізку $[0,1]$.

Властивість 2. *Функція* $F(x)$ *є неспадною, тобто, якщо* $x_1 < x_2$, *то* $F(x_1) \le F(x_2)$.

Доведення. Розглянемо:

$$F(x_2) = P(X < x_2) = P(X < x_1) + P(x_1 \le X < x_2) \ge P(X < x_1) = F(x_1),$$

оскільки $P(x_1 \le X < x_2) \ge 0$ (будь-яка ймовірність є невід'ємним числом).

Водночас отримуємо рівність:

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1). \quad (*)$$

Якщо випадкова величина X неперервна, то можна показати, що її функція розподілу також неперервна, тобто

$$\lim_{\Delta x \to 0} F(x + \Delta x) = F(x).$$

Цей факт можна проілюструвати на підставі рівності (*). Покладемо у цій рівності $x_1 = x$, $x_2 = x + \Delta x$. Дістанемо:

$$P(x \le X < x + \Delta x) = F(x + \Delta x) - F(x). \quad (**)$$

Ймовірність у лівій частині цієї рівності є ймовірність потрапляння випадкової точки (значення НВВ X) в інтервал, довжина якого дорівнює Δx. І якщо Δx прямує до нуля, то на підставі формули геометричної ймовірності ймовірність $P(x \le X < x + \Delta x)$ також прямує до нуля, отже:

$$\lim_{\Delta x \to 0} (F(x + \Delta x) - F(x)) = 0,$$

а це й означає неперервність функції $F(x)$.

Властивість 3. *Ймовірність того, що НВВ X прийме одне певне значення, дорівнює нулю.*

Ця властивість безпосередньо випливає з рівності:

$$P(X = x) = \lim_{\Delta x \to 0} P(x \le X < x + \Delta x) = \lim_{\Delta x \to 0} (F(x + \Delta x) - F(x)) = 0\,.$$

Наслідок. *Для НВВ X мають місце рівності:*

$$P(a \le X \le b) = P(a < X < b) = P(a < X \le b) = P(a \le X \le b) = F(b) - F(a)\,.$$

Дійсно, розглянемо, наприклад:

$$P(a < X \le b) = P(a < X < b) + P(X = b) = P(a < X < b) + 0 = P(a < X < b)\,.$$

Властивість 3 цікава, і може мати невірне тлумачення. З цієї властивості, здавалось би, випливає, що подія $X = x$ неможлива, адже її ймовірність дорівнює нулю. І внаслідок довільності x звідси випливав би парадоксальний висновок, що НВВ X не може прийняти ніякого певного значення. В той же час очевидно, що внаслідок випробування випадкова величина X обов'язково прийме деяке певне значення. Отже, ми бачимо, що з того, що ймовірність події дорівнює нулю, не випливає, взагалі кажучи, що ця подія неможлива. Глибокі причини цього факту розкриваються в більш детальних курсах теорії ймовірностей, які ґрунтуються на теорії міри та теорії функцій дійсного змінного.

Властивість 4. *Якщо всі можливі значення випадкової величини належать інтервалу (a,b), то $F(x) = 0$ при $x \le a$ і $F(x) = 1$ при $x \ge b$.*

Доведення. Нехай $x_1 \le a$. Тоді подія $X < x_1$ неможлива (значень, менших x_1, величина X не приймає), отже $F(x_1) = P(X < x_2) = 0$.

Нехай $x_2 \ge b$. Тоді подія $X < x_2$ достовірна (всі можливі значення величини X менше x_2), отже $F(x_2) = P(X < x_2) = 1$.

Властивість 5. *Справедливі наступні співвідношення:*

$$\lim_{x \to -\infty} F(x) = 0\,, \quad \lim_{x \to +\infty} F(x) = 1\,.$$

Доведення. Дійсно, розглянемо:

$$\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(X < x) = P(X < -\infty) = 0$$

(значень, менших $-\infty$, зрозуміло, не існує).

$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X < x) = P(X < +\infty) = 1$$

(всі можливі значення випадкової величини менше $+\infty$).

Приклади

1. Функція розподілу НВВ X задається рівністю:

$$F(x) = A + B\operatorname{arctg} x\,, \quad -\infty < x < +\infty\,.$$

Визначити сталі A і B і знайти $P(0 < X < 1)$.

Маємо:

$$\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty}(A + B\operatorname{arctg} x) = A + B\lim_{x\to-\infty}\operatorname{arctg} x = A - B\frac{\pi}{2} = 0,$$

$$\lim_{x\to+\infty} F(x) = \lim_{x\to+\infty}(A + B\operatorname{arctg} x) = A + B\lim_{x\to+\infty}\operatorname{arctg} x = A + B\frac{\pi}{2} = 1.$$

Отже для сталих A і B отримали систему двох лінійних алгебраїчних рівнянь:

$$\begin{cases} A - B\dfrac{\pi}{2} = 0, \\ A + B\dfrac{\pi}{2} = 1. \end{cases}$$

Розв'язуючи цю систему, отримуємо:

$$A = \frac{1}{2},\ B = \frac{1}{\pi}.$$

Отже:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x,\ -\infty < x < +\infty.$$

Графік цієї функції зображено на рис. 19.

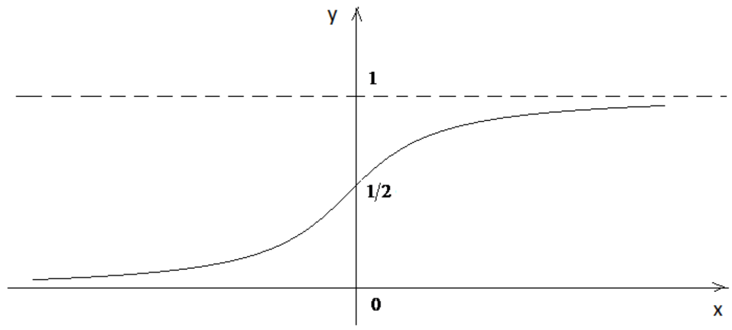


Рис. 19

Знайдемо тепер:

$$P(0 < X < 1) = F(1) - F(0) = \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}1\right) - \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}0\right) =$$

$$= \frac{1}{\pi}\operatorname{arctg}1 = \frac{1}{\pi}\cdot\frac{\pi}{4} = \frac{1}{4}.$$

2. НВВ X задана своєю функцією розподілу:

$$F(x) = \begin{cases} 0,\ x \le 0, \\ A + B\sin x,\ 0 < x \le \pi/4, \\ 1,\ x > \pi/4. \end{cases}$$

Визначити сталі A і B і знайти ймовірність того, що внаслідок 4-х незалежних випробувань величина X рівно 3 рази прийме значення, яке належить інтервалу $(\pi/6, \pi/3)$.

Сталі A і B визначимо з умови неперервності функції $F(x)$, зокрема, у точках $x=0$ та $x=\pi/4$. Границі справа та зліва функції $F(x)$ в цих точках мають збігатися, отже:

$$A+B\sin 0=0,\ A+B\sin\frac{\pi}{4}=1.$$

Таким чином, отримали систему лінійних алгебраїчних рівнянь:

$$\begin{cases} A=0, \\ A+B\dfrac{1}{\sqrt{2}}=1. \end{cases}$$

Звідси отримуємо: $A=0,\ B=\sqrt{2}$. Отже:

$$F(x)=\begin{cases} 0,\ x\le 0, \\ \sqrt{2}\sin x,\ 0<x\le \pi/4, \\ 1,\ x>\pi/4. \end{cases}$$

Знайдемо тепер:

$$p=P\left(\frac{\pi}{6}<X<\frac{\pi}{3}\right)=F\left(\frac{\pi}{3}\right)-F\left(\frac{\pi}{6}\right)=1-\sqrt{2}\sin\frac{\pi}{6}=1-\sqrt{2}\cdot\frac{1}{2}=\frac{\sqrt{2}-1}{\sqrt{2}}.$$

Ймовірність того, що внаслідок 4-х випробувань значення з інтервалу $(\pi/6,\ \pi/3)$ випадкова величина X прийме рівно 3 рази, знайдемо за формулою Бернуллі (див. п. 1.13):

$$P_4(3)=C_4^3p^3(1-p)=4\cdot\left(\frac{\sqrt{2}-1}{\sqrt{2}}\right)^3\frac{1}{\sqrt{2}}=\frac{4\left(\sqrt{2}-1\right)^3}{4}=\left(\sqrt{2}-1\right)^3\approx 0{,}071.$$

2.16. Функція розподілу дискретної випадкової величини

Функцію розподілу можна будувати не тільки для неперервної випадкової величини, але й для дискретної.

Приклад 1. Побудувати функцію розподілу ДВВ X, яку задано законом розподілу:

X	3	4	7	10
P	0,2	0,1	0,4	0,3

Розіб'ємо числову пряму на інтервали, межі яких визначаються значеннями ДВВ X, і побудуємо функцію розподілу на кожному інтервалі окремо.

а) $-\infty < x \le 3$: $F(x) = P(X < x) = P(X < x \le 3) = 0$ (адже значень, менших 3, величина X не приймає);

б) $3 < x \le 4$: $F(x) = P(X < x) = P(X = 3) = 0,2$ (адже у цьому випадку є лише єдине значення величини X, яке менше, ніж x, це $X = 3$);

в) $4 < x \le 7$: $F(x) = P(X < x) = P(X = 3) + (X = 4)) = P(X = 3) + P(X = 4) =$ $= 0,2 + 0,1 = 0,3$ (адже у цьому випадку є два значення величини X, які менші, ніж x, це $X = 3$ та $X = 4$);

г) $7 < x \le 10$: $F(x) = P(X < x) = P(X = 3) + (X = 4) + (X = 7)) =$ $= P(X = 3) + P(X = 4) + P(X = 7) = 0,2 + 0,1 + 0,4 = 0,7$ (три значення величини X, які менші, ніж x, це $X = 3$, $X = 4$, $X = 7$);

д) $10 < x < +\infty$: $F(x) = P(X < x) = 1$ (тепер всі значення величини X є меншими, ніж x).

Графік функції зображено на рис. 20.

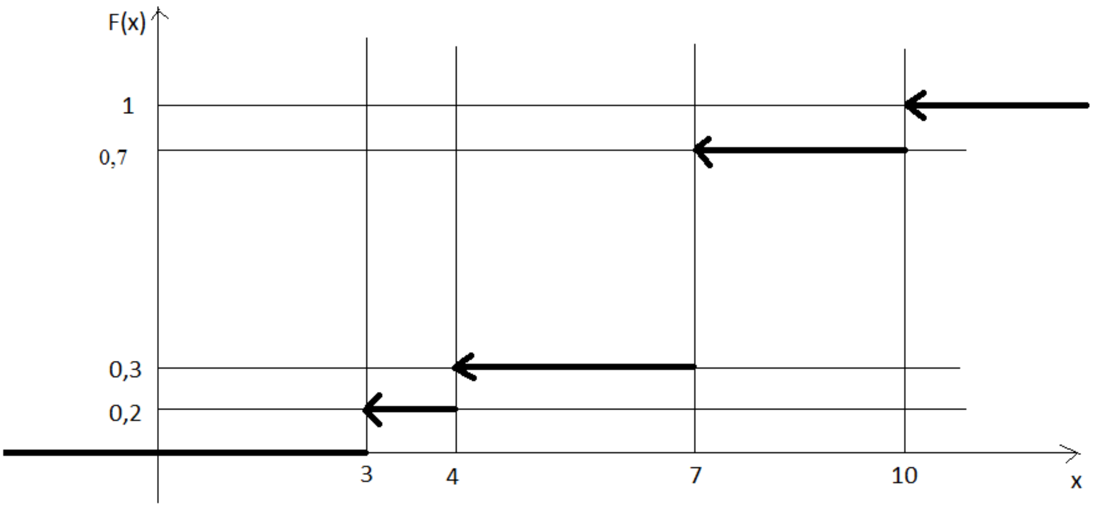


Рис. 20

Цей графік характерний взагалі для функції розподілу ДВВ. На відміну від функції розподілу НВВ, для ДВВ функція розподілу не є неперервною: вона кусково-неперервна та кусково-стала. В точках, які відповідають значенням випадкової величини, ця функція є неперервною зліва, а справа – розрив I роду (це показується стрілками). Для знаходження значень функції розподілу на кожному інтервалі, очевидно, треба знайти суму ймовірностей значень випадкової величини, що лежать зліва від цього інтервалу.

Приклад 2. Монета кидається до першої появи герба. Побудувати функцію розподілу числа кидань.

Число кидань монети розподілено за геометричним законом (див. п. 2.13):

X	1	2	3	$\dots$	k	$k+1$	$\dots$
P	$\frac{1}{2}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$	$\dots$	$\frac{1}{2^k}$	$\frac{1}{2^{k+1}}$	$\dots$

Розбиваємо числову пряму на інтервали, що визначаються значеннями випадкової величини (цього разу кількість інтервалів буде нескінченною), і будуємо функцію розподілу на кожному інтервалі окремо:

$-\infty < x \le 1$: $F(x) = 0$;

$1 < x \le 2$: $F(x) = \frac{1}{2}$;

$2 < x \le 3$: $F(x) = \frac{1}{2} + \frac{1}{2^2}$;

$3 < x \le 4$: $F(x) = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3}$;

$\dots$

$$k < x \le k+1: \quad F(x) = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^k} = \frac{\frac{1}{2}\left(1 - \frac{1}{2^k}\right)}{1 - \frac{1}{2}} = 1 - \frac{1}{2^k};$$

$\dots$

2.17. Щільність розподілу неперервної випадкової величини

Нехай задано НВВ X, і $F(x)$ – її функція розподілу.

Означення. *Щільністю розподілу* НВВ X називається функція

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

Таким чином, функція $F(x)$ є первісною для функції $f(x)$.

Приклад. Нехай функція розподілу НВВ X має вигляд:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x, \quad -\infty < x < +\infty.$$

Знайти щільність розподілу.

Маємо:

$$f(x)=F'(x)=\left(\frac{1}{2}+\frac{1}{\pi}\operatorname{arctg}x\right)'=\frac{1}{\pi\left(1+x^2\right)},\ -\infty<x<+\infty .$$

Теорема. *Якщо* X *– НВВ, і* $f(x)$ *– її щільність розподілу, то*

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

Доведення. Використовуючи рівність (*) (п. 2.15) і наслідок з Властивості 3 функції розподілу, маємо:

$$P(a<X<b)=F(b)-F(a).$$

За формулою Ньютона – Лейбніца:

$$F(b)-F(a)=\int_a^b f(x)dx ,$$

звідки й випливає твердження теореми.

З геометричної точки зору це означає, що ймовірність того, що НВВ X внаслідок випробування прийме значення, що належить інтервалу (a,b), дорівнює площі S криволінійної трапеції, яка обмежена віссю Ox, графіком функції $y=f(x)$ і прямими $x=a$, $x=b$ (рис. 21).

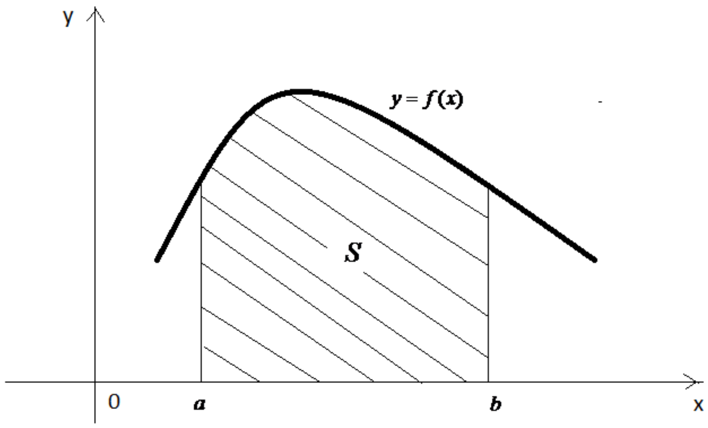


Рис. 21

$$S=P(a<X<b).$$

Висота графіка щільності на певних ділянках числової прямої показує густину групування значень випадкової величини на цих ділянках. Найбільша густина буде там, де більша висота графіка (тобто більше величина щільності розподілу, більше ймовірність того, що внаслідок випробування випадкова величина прийме значення саме з цієї ділянки), а де висота графіка мала, там й густина значень також незначна. На тих ділянках, де щільність нульова, взагалі нема можливих значень випадкової величини.

<u>Приклад</u>. Задано щільність розподілу НВВ X :

$$f(x)=\begin{cases} 0,\ x\le 0,\\ \cos x,\ 0<x\le \pi/2,\\ 0,\ x>\pi/2. \end{cases}$$

Знайти $P(\pi/4<X<\pi)$.

Згідно з теоремою маємо:

$$P\left(\frac{\pi}{4}<X<\pi\right)=\int_{\pi/4}^{\pi} f(x)dx=\int_{\pi/4}^{\pi/2} f(x)dx+\int_{\pi/2}^{\pi} f(x)dx=$$

$$=\int_{\pi/4}^{\pi/2}\cos x\,dx+\int_{\pi/2}^{\pi}0\cdot dx=\sin x\Big|_{\pi/4}^{\pi/2}=\sin\frac{\pi}{2}-\sin\frac{\pi}{4}=1-\frac{1}{\sqrt{2}}.$$

Наслідок. *Якщо* $F(x)$ *– функція розподілу НВВ* X*, а* $f(x)$ *– щільність розподілу цієї величини, то*

$$F(x)=\int_{-\infty}^{x} f(t)\,dt.$$

Доведення. За означенням функції розподілу маємо:

$$F(x)=P(X<x)=P(-\infty<X<x)=\int_{-\infty}^{x} f(t)\,dt.$$

Внаслідок рівності $f(x)=F'(x)$ і цього Наслідку, функцію $f(x)$ іноді називають *диференціальною функцією розподілу*, а функцію $F(x)$ *– інтегральною функцією розподілу.*

<u>Приклад</u>. Задано щільність розподілу НВВ X:

$$f(x)=\begin{cases} 0,\ x\le 1,\\ x-1/2,\ 1<x\le 2,\\ 0,\ x>2. \end{cases}$$

Знайти інтегральну функцію розподілу.

Оскільки щільність розподілу задано різними виразами на різних інтервалах, то й функцію розподілу шукатимемо на кожному інтервалі окремо.

1). $-\infty<x\le 1$:

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\int_{-\infty}^{x}0\cdot dt=0.$$

2). $1<x\le 2$:

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\int_{-\infty}^{1} f(t)\,dt+\int_{-1}^{x} f(t)\,dt=\int_{-\infty}^{1} 0\cdot dt+\int_{-1}^{x}\left(t-\frac{1}{2}\right)dt=$$

$$=\left.\frac{(t-1/2)^2}{2}\right|_{1}^{x}=\frac{(x-1/2)^2}{2}-\frac{(1-1/2)^2}{2}=\frac{x^2-x}{2}.$$

3). $x>2$:

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\int_{-\infty}^{1} f(t)\,dt+\int_{1}^{2} f(t)\,dt+\int_{2}^{x} f(t)\,dt=$$

$$=\int_{-\infty}^{1} 0\cdot dt+\int_{1}^{2}\left(t-\frac{1}{2}\right)dt+\int_{2}^{x} 0\cdot dt=\int_{1}^{2}\left(t-\frac{1}{2}\right)dt=\left.\frac{(t-1/2)^2}{2}\right|_{1}^{2}=$$

$$=\frac{(2-1/2)^2}{2}-\frac{(1-1/2)^2}{2}=\frac{9}{8}-\frac{1}{8}=1.$$

Отже:

$$F(x)=\begin{cases} 0, & x\le 1,\\ \dfrac{x^2-x}{2}, & 1<x\le 2,\\ 1, & x>2.\end{cases}$$

Ці обчислення можна спростити, якщо скористатися тим, що щільність на 1-му та 3-му інтервалах дорівнює нулю, отже, випадкова величина не приймає значення з цих інтервалів, а тоді, згідно з Властивістю 4 функції розподілу (див. п. 2.15), на 1-му інтервалі функція $F(x)$ дорівнює нулю, а на 3-му – одиниці.

2.18. Властивості щільності розподілу

Властивість 1. *Щільність розподілу є функція невід'ємна:*

$f(x)\ge 0$.

Доведення. Функція розподілу $F(x)$ є функція неспадна, отже її похідна $F'(x)=f(x)$ – функція невід'ємна.

Властивість 2. *Невласний інтеграл від щільності розподілу у межах від* $-\infty$ *до* $+\infty$ *дорівнює одиниці:*

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

Доведення. Маємо:

$$\int_{-\infty}^{+\infty} f(x)\,dx = P(-\infty < X < +\infty) = 1.$$

Рівність (*) називають *умовою нормування* для НВВ.

Приклад. Задано щільність розподілу НВВ:

$$f(x) = \frac{A}{4+x^2},\ -\infty < x < +\infty.$$

Знайти сталу A.

Маємо:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{+\infty} \frac{A}{4+x^2}dx = A\int_{-\infty}^{+\infty} \frac{dx}{4+x^2} = 2A\int_{-0}^{+\infty} \frac{dx}{4+x^2} =$$

$$= 2A\lim_{b\to+\infty}\int_0^b \frac{dx}{4+x^2} = 2A\lim_{b\to+\infty}\left(\frac{1}{2}\operatorname{arctg}\frac{x}{2}\Big|_0^b\right) = 2A\lim_{b\to+\infty}\frac{1}{2}\operatorname{arctg}\frac{b}{2} = 2A\cdot\frac{\pi}{4} = 1.$$

Звідси $A = 2/\pi$,

$$f(x) = \frac{2}{\pi\left(4+x^2\right)}.$$

Властивість 3. *Якщо всі можливі значення НВВ X належать відрізку $[a,b]$, то $f(x)$ тотожно дорівнює нулю при $x<a$ та при $x>b$, причому:*

$$\int_a^b f(x)\,dx = 1.$$

Доведення. Користуючись Властивістю 4 функції розподілу (див. п. 2.15), маємо, що $F(x)\equiv 0$ при $x<a$, і $F(x)\equiv 1$ при $x>b$. Звідси й випливає, що $f(x) = F'(x) \equiv 0$ при $x<a$ і при $x>b$, адже похідна сталої дорівнює нулю. А тоді з умови нормування отримаємо:

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

$$= \int_{-\infty}^{a} 0\cdot dx + \int_a^b f(x)\,dx + \int_b^{+\infty} 0\cdot dx = \int_a^b f(x)\,dx = 1.$$

Приклад. Задано щільність розподілу НВВ X:

$$f(x) = \begin{cases} 0, & x<1, \\ A\ln x, & 1\le x\le e^2, \\ 0, & x>e^2. \end{cases}$$

найти параметр A.

Оскільки функція $f(x)$ відмінна від нуля лише на відрізку $[1,e^2]$, то

$$\int_1^{e^2} f(x)\,dx = 1.$$

Звідси маємо:

$$\int_1^{e^2} f(x)\,dx = \int_1^{e^2} A\ln x\,dx = A\int_1^{e^2}\ln x\,dx \left[\begin{matrix} u=\ln x, & dx=dx, \\ du=\dfrac{dx}{x}, & v=x \end{matrix}\right] =$$

$$= A\left(x\ln x\Big|_1^{e^2} - x\Big|_1^{e^2}\right) = A(e^2\ln e^2 - e^2 + 1) = A(2e^2 - e^2 + 1) = A(e^2+1) = 1.$$

Отже:

$$A = \frac{1}{e^2+1},$$

$$f(x) = \begin{cases} 0, & x<1, \\ \dfrac{1}{e^2+1}\ln x, & 1\le x\le e^2, \\ 0, & x>e^2. \end{cases}$$

2.19. Ймовірнісний зміст щільності розподілу

Нехай $F(x)$ – функція розподілу НВВ X. За означенням щільності розподілу:

$$f(x) = F'(x) = \lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}.$$

Оскільки $F(x+\Delta x)-F(x) = P(x<X<x+\Delta x)$, то

$$f(x) = \lim_{\Delta x\to 0}\frac{P(x<X<x+\Delta x)}{\Delta x}.$$

Якщо ми цю границю наближено замінимо виразом, що стоїть під знаком границі, то дістанемо:

$$f(x) \approx \frac{P(x<X<x+\Delta x)}{\Delta x}.$$

Ця наближена рівність буде тим точніша, чим менше Δx. Таким чином, можна стверджувати, що щільність розподілу наближено дорівнює відношенню ймовірності потрапляння НВВ X у малий інтервал до довжини цього інтервалу.

2.20. Числові характеристики неперервних випадкових величин

Нехай X – НВВ, і $f(x)$ – її щільність розподілу.

Означення. *Математичним сподіванням* НВВ X називається число

$$M[X]=\int_{-\infty}^{+\infty} x\,f(x)\,dx,$$

за умови, що цей інтеграл збігається абсолютно, тобто існує інтеграл

$$\int_{-\infty}^{+\infty} |x|\,f(x)\,dx.$$

Якщо, зокрема, всі можливі значення НВВ X належать відрізку $[a,b]$, то

$$M[X]=\int_a^b x\,f(x)\,dx.$$

Легко зрозуміти при цьому, що тоді $M[X]$ також належить відрізку $[a,b]$. Дійсно, оскільки $a\le x\le b$ і $f(x)\ge 0$, то

$$\int_a^b a\,f(x)\,dx\le\int_a^b x\,f(x)\,dx\le\int_a^b b\,f(x)\,dx.$$

Або

$$a\int_a^b f(x)\,dx\le M[X]\le b\int_a^b f(x)\,dx.$$

А оскільки

$$\int_a^b f(x)\,dx=1,$$

то

$$a\le M[X]\le b.$$

Можна довести, що для НВВ математичне сподівання має такі ж властивості, як і для ДВВ (див. п. 2.5):

1) $M[C]=C$;
2) $M[CX]=CM[X]$;
3) для незалежних НВВ X і Y: $M[XY]=M[X]M[Y]$;
4) $M[X\pm Y]=M[X]\pm M[Y]$.

Означення. *Дисперсією* НВВ X називається число

$$D[X]=\int_{-\infty}^{+\infty}(x-m)^2 f(x)\,dx\,,$$

де $m=M[X]$, за умови, що цей інтеграл збігається. Якщо, зокрема, всі значення НВВ X належать відрізку $[a,b]$, то

$$D[X]=\int_{a}^{b}(x-m)^2 f(x)\,dx\,.$$

Дисперсія НВВ також зберігає властивості дисперсії ДВВ, а саме:

1) $D[X]\geq 0$;
2) $D[C]=0$;
3) $D[CX]=C^2D[X]$;
4) для незалежних НВВ X і Y: $D[X\pm Y]=D[X]+D[Y]$.

Також справедлива наступна формула:

$$D[X]=\int_{-\infty}^{+\infty}x^2 f(x)\,dx-m^2\,,$$

тобто $D[X]=M[X^2]-M^2[X]$.

Означення. *Середньоквадратичним відхиленням* НВВ X називається арифметичний квадратний корінь з її дисперсії:

$$\sigma[X]=\sqrt{D[X]}\,.$$

Приклад. НВВ X задано її щільністю розподілу:

$$f(x)=\begin{cases}\frac{1}{2}x, & x\in(0,2),\\ 0, & x\notin(0,2).\end{cases}$$

Знайти $M[X]$, $D[X]$, $\sigma[X]$.

Оскільки $f(x)$ відмінна від нуля лише на інтервалі $(0,2)$, тобто всі можливі значення величини X належать цьому інтервалу, то

$$m=M[X]=\int_0^2 x\,f(x)\,dx=\int_0^2 x\cdot\frac{1}{2}x\,dx=\frac{1}{2}\int_0^2 x^2\,dx=\left.\frac{x^3}{6}\right|_0^2=\frac{8}{6}=\frac{3}{4};$$

$$D[X]=\int_0^2 x^2 f(x)\,dx-m^2=\int_0^2 x^2\cdot\frac{1}{2}x\,dx-\left(\frac{3}{4}\right)^2=\frac{1}{2}\int_0^2 x^3\,dx-\frac{9}{16}=$$

$$=\left.\frac{x^4}{8}\right|_0^2-\frac{9}{16}=2-\frac{9}{16}=\frac{23}{16}\;;$$

$$\sigma[X]=\sqrt{D[X]}=\sqrt{\frac{23}{16}}=\frac{\sqrt{23}}{4}\approx 1{,}2\,.$$

2.21. Рівномірний розподіл

При розв’язанні практичних задач доводиться мати справу з різними розподілами НВВ. Щільності розподілів називають також *законами розподілів*. Тут ми розглянемо рівномірний розподіл.

Означення. НВВ X називається *розподіленою рівномірно* на відрізку $[a, b]$, якщо її щільність розподілу має вигляд:

$$f(x)=\begin{cases} h, \ x \in [a, b], \\ 0, \ x \notin [a, b]. \end{cases}$$

Тут h – стала. Рівномірно розподілені величини зустрічаються досить часто, це величини, які з однаковою ймовірністю можуть прийняти будь-яке значення з деякого відрізку. Наприклад, під час контрольного опиту на уроці учень може бути викликаний вчителем у будь-яку мить уроку. Можна сказати, що момент часу виклику певного учня є НВВ, рівномірно розподіленою на відрізку [0, 45] (45 хвилин продовжується урок).

Оскільки щільність розподілу відмінна від нуля лише на відрізку $[a, b]$, то наша випадкова величина може приймати свої значення лише з цього відрізку. Визначимо сталу h, для чого скористаємось Властивістю 3 щільності розподілу:

$$\int_a^b h\,dx = 1 .$$

Звідси маємо:

$$\int_a^b h\,dx = h\int_a^b dx = h(b-a) = 1 ,$$

отже:

$$h=\frac{1}{b-a} .$$

Таким чином, уточнена щільність розподілу має вигляд:

$$f(x)=\begin{cases} \dfrac{1}{b-a}, \ x \in [a, b], \\ 0, \quad x \notin [a, b]. \end{cases}$$

Графік щільності наведено на рис. 22.

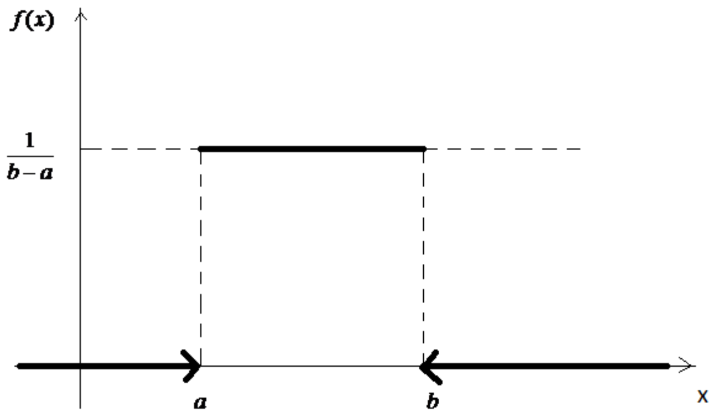


Рис. 22

Побудуємо тепер функцію розподілу. Оскільки наша випадкова величина приймає свої значення лише на відрізку $[a,b]$, то, користуючись Властивістю 4 функції розподілу (див. п. 2.15), можна записати, що $F(x)=0$ при $x\le a$, і $F(x)=1$ при $x\ge b$. Знайдемо $F(x)$ при $a<x<b$:

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\int_{-\infty}^{a} f(t)\,dt+\int_{a}^{x} f(t)\,dt=\int_{-\infty}^{a} 0\cdot dt+\int_{a}^{x}\frac{1}{b-a}dt=$$

$$=\frac{1}{b-a}\int_{a}^{x} dt=\frac{x-a}{b-a}.$$

Отже:

$$F(x)=\begin{cases} 0, & x\le a, \\ \dfrac{x-a}{b-a}, & a<x<b, \\ 1, & x\ge b. \end{cases}$$

Графік функції $F(x)$ наведено на рис. 23.

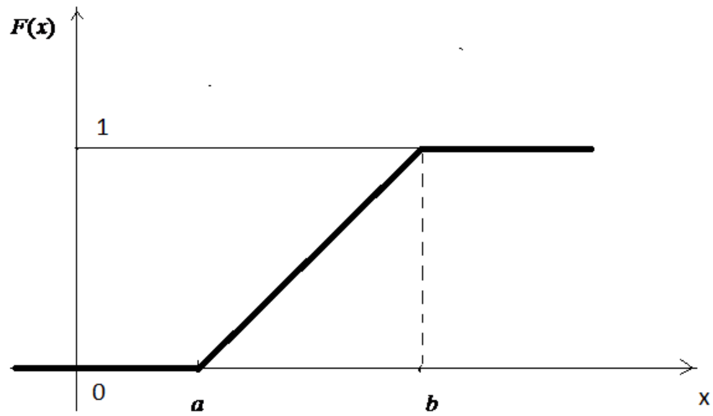


Рис. 23

Знайдемо тепер числові характеристики рівномірно розподіленої величини X. Почнемо з математичного сподівання.

$$M[X]=\int_a^b x\,f(x)\,dx=\int_a^b x\frac{1}{b-a}\,dx=\frac{1}{b-a}\int_a^b x\,dx=\frac{1}{b-a}\cdot\frac{x^2}{2}\bigg|_a^b=$$

$$=\frac{b^2-a^2}{2(b-a)}=\frac{(b-a)(b+a)}{2(b-a)}=\frac{a+b}{2}.$$

Тобто математичне сподівання рівномірного розподілу – це середня точка відрізку $[a,b]$. Взагалі, якщо розподіл такий, що графік його щільності симетричний відносно прямої $x=x_0$, то математичне сподівання розподілу дорівнює x_0.

Знайдемо дисперсію:

$$D[X]=\int_a^b x^2 f(x)\,dx-\left(\frac{a+b}{2}\right)^2=\int_a^b x^2\cdot\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2=$$

$$=\frac{1}{b-a}\int_a^b x^2\cdot dx-\left(\frac{a+b}{2}\right)^2=\frac{1}{b-a}\cdot\frac{x^3}{3}\bigg|_a^b-\left(\frac{a+b}{2}\right)^2=\frac{b^3-a^3}{3(b-a)}-\frac{(a+b)^2}{4}=$$

$$=\frac{a^2+ab+b^2}{3}-\frac{a^2+2ab+b^2}{4}=\frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12}=$$

$$=\frac{b^2-2ab+a^2}{12}=\frac{(b-a)^2}{12}.$$

Тепер середньоквадратичне відхилення:

$$\sigma[X]=\sqrt{\frac{(b-a)^2}{12}}=\frac{b-a}{2\sqrt{3}}.$$

Приклад. НВВ X розподілена рівномірно на відрізку $[2,8]$. Побудувати функцію розподілу, знайти математичне сподівання, дисперсію, середньоквадратичне відхилення.

Щільність розподілу має вигляд:

$$f(x)=\begin{cases}\frac{1}{6},\ x\in[2,8],\\ 0,\ x\notin[2,8].\end{cases}$$

Функція розподілу:

$$F(x)=\begin{cases}0,\ x\le 2,\\ \frac{x-2}{6},\ 2<x<8,\\ 1,\ x\ge 8.\end{cases}$$

За формулами, що отримано вище, маємо:

$$M[X]=5,\ D[X]=3,\ \sigma[X]=\sqrt{3}\,.$$

2.22. Показниковий розподіл

Означення. НВВ X називається розподіленою за *показниковим (експоненціальним)* законом з параметром $\lambda>0$, якщо її щільність розподілу має вигляд:

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

Графік щільності наведено на рис. 24.

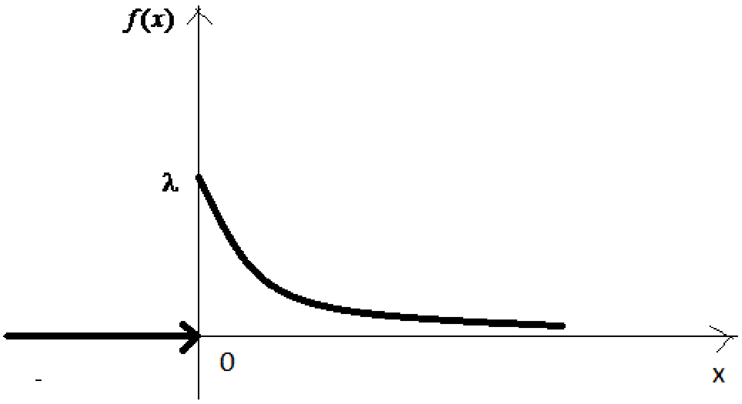


Рис. 24

Така щільність показує, що величина X може приймати лише значення з півінтервалу $[0,+\infty)$, причому ці значення групуються головним чином біля початку координат, а при віддаленні від нього ймовірність появи цих значень швидко спадає.

Показниковий розподіл часто використовується для опису інтервалів між послідовними рідкими подіями, наприклад, між відвідуваннями непопулярних сайтів.

Перевіримо виконання умови нормування. Знайдемо:

$$\int_{-\infty}^{+\infty} f(x)\,dx=\int_{-\infty}^{0} f(x)\,dx+\int_{0}^{+\infty} f(x)\,dx=\int_{-\infty}^{0} 0\cdot dx+\int_{0}^{+\infty}\lambda e^{-\lambda x}\,dx=\lim_{b\to+\infty}\int_{0}^{b}\lambda e^{-\lambda x}\,dx=$$

$$=-\lim_{b\to+\infty}\int_{0}^{b} e^{-\lambda x}\,d(-\lambda x)=-\lim_{b\to+\infty}\left(e^{-\lambda x}\Big|_{0}^{b}\right)=-\lim_{b\to+\infty}(e^{-\lambda b}-1)=1$$

(на підставі того, що $\lim\limits_{b\to+\infty} e^{-\lambda b}=0$).

Знайдемо функцію розподілу. При $x<0$ маємо:

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{x} 0\cdot dt=0.$$

При $x \ge 0$:

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{0} f(t)dt+\int_{0}^{x} f(t)dt=\int_{-\infty}^{0} 0\cdot dt+\int_{0}^{x}\lambda e^{-\lambda t}dt=$$

$$=-\int_{0}^{x} e^{-\lambda t}d(-\lambda t)=-e^{-\lambda t}\Big|_0^x=1-e^{-\lambda x}.$$

Отже:

$$F(x)=\begin{cases} 0, & x<0, \\ 1-e^{-\lambda x}, & x\ge 0. \end{cases}$$

Графік функції $F(x)$ наведено на рис. 25.

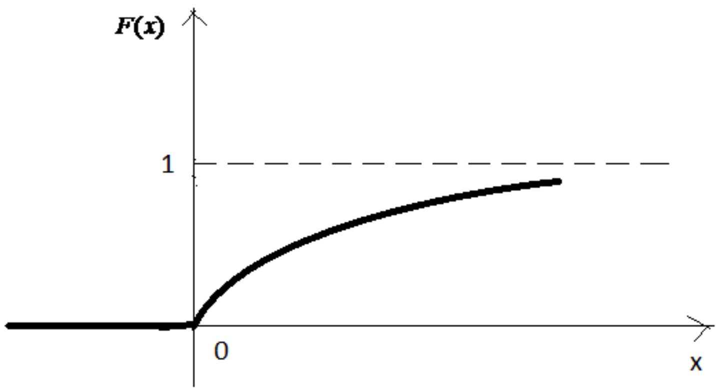


Рис. 25

Знаючи $F(x)$, знайдемо ймовірність потрапляння НВВ X в інтервал (a, b) (припускаємо, що $0<a<b$):

$$P(a<X<b)=F(b)-F(a)=1-e^{-\lambda b}-(1-e^{-\lambda a})=e^{-\lambda a}-e^{-\lambda b}. \qquad (*)$$

<u>Приклад.</u> НВВ X розподілена за показниковим законом з щільністю:

$$f(x)=\begin{cases} 0, & x<0, \\ 2e^{-2x}, & x\ge 0. \end{cases}$$

Знайти ймовірність того, що внаслідок випробування величина X потрапить в інтервал $(0,4;\ 0,6)$.

За умовою $\lambda=2$. За формулою (*):

$$P(0,4<X<0,6)=e^{-2\cdot 0,4}-e^{-2\cdot 0,6}=e^{-0,8}-e^{-1,2}\approx 0,44933-0,30119\approx 0,148.$$

Знайдемо числові характеристики показникового розподілу. Почнемо з математичного сподівання. Інтегруючи за частинами, дістанемо:

$$M[X]=\int\limits_{-\infty}^{+\infty} x\,f(x)\,dx=\int\limits_{0}^{+\infty} x\lambda e^{-\lambda x}\,dx=-\int\limits_{0}^{+\infty} x\,d(e^{-\lambda x})=-\lim_{b\to+\infty}\int\limits_{0}^{b} x\,d(e^{-\lambda x})=$$

$$=-\lim_{b\to+\infty}\left(xe^{-\lambda x}\Big|_0^b-\int\limits_0^b e^{-\lambda x}dx\right)=-\lim_{b\to+\infty}\left(be^{-\lambda b}+\frac{1}{\lambda}e^{-\lambda x}\Big|_0^b\right)=-\lim_{b\to+\infty}\left(\frac{b}{e^{\lambda b}}+\frac{1}{\lambda}(e^{-\lambda b}-1)\right).$$

Оскільки $\lim\limits_{b\to+\infty} e^{-\lambda b}=0$, і за правилом Лопіталя

$$\lim_{b\to+\infty}\frac{b}{e^{\lambda b}}=\lim_{b\to+\infty}\frac{1}{\lambda e^{\lambda b}}=0,$$

то

$$M[X]=\frac{1}{\lambda}.$$

Таким чином, математичне сподівання показникового розподілу дорівнює величині, оберненій параметру λ.

Знайдемо дисперсію за формулою:

$$D[X]=\int\limits_0^{+\infty} x^2\lambda e^{-\lambda x}dx-\frac{1}{\lambda^2}.$$

Інтегруючи за частинами, дістанемо:

$$\int\limits_0^{+\infty} x^2\lambda e^{-\lambda x}dx=\lim_{b\to+\infty}\int\limits_0^b x^2\lambda e^{-\lambda x}dx=-\lim_{b\to+\infty}\int\limits_0^b x^2 d(e^{-\lambda x})=$$

$$=-\lim_{b\to+\infty}\left(x^2e^{-\lambda x}\Big|_0^b-2\int\limits_0^b xe^{-\lambda x}dx\right)=-\lim_{b\to+\infty}\left(\frac{b^2}{e^{\lambda b}}+\frac{2}{\lambda}\int\limits_0^b x\,d(e^{-\lambda x})\right)=$$

$$=-\lim_{b\to+\infty}\left(\frac{b^2}{e^{\lambda b}}+\frac{2}{\lambda}\left(xe^{-\lambda x}\Big|_0^b-\int\limits_0^b e^{-\lambda x}dx\right)\right)=-\lim_{b\to+\infty}\left(\frac{b^2}{e^{\lambda b}}+\frac{2}{\lambda}\left(\frac{b}{e^{\lambda b}}+\frac{1}{\lambda}(e^{-\lambda b}-1)\right)\right)=$$

$$=-\lim_{b\to+\infty}\left(\frac{b^2}{e^{\lambda b}}+\frac{2b}{\lambda e^{\lambda b}}+\frac{2}{\lambda^2}e^{-\lambda b}+\frac{2}{\lambda^2}\right).$$

Оскільки

$$\lim_{b\to+\infty}\frac{b^2}{e^{\lambda b}}=\lim_{b\to+\infty}\frac{2b}{\lambda e^{\lambda b}}=\lim_{b\to+\infty}e^{-\lambda b}=0,$$

то

$$\int\limits_0^{+\infty} x^2\lambda e^{-\lambda x}dx=\frac{2}{\lambda^2}.$$

А тоді

$$D[X]=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

Тепер знайдемо середньоквадратичне відхилення:

$$\sigma[X]=\sqrt{D[X]}=\sqrt{\frac{1}{\lambda^2}}=\frac{1}{\lambda},$$

тобто математичне сподівання та середньоквадратичне відхилення показникового розподілу співпадають і обернені за величиною параметру λ.

Приклад. Знайти дисперсію та середньоквадратичне відхилення показникового розподілу, яке задано інтегральною функцією:

$$F(x)=1-e^{-0,4x} \quad (x\geq 0).$$

Параметр λ дорівнює 0,4, тому

$$M[X]=\frac{1}{0,4}=2,5;$$

$$D[X]=\frac{1}{0,4^2}=\frac{1}{0,16}=6,25;$$

$$\sigma[X]=\sqrt{D[X]}=\sqrt{6,25}=2,5.$$

2.23. Показниковий закон надійності

Нехай є деякий технічний елемент (наприклад, комп'ютер), який працює на протязі деякого проміжку часу. Але за цей час може відбутися відмова цього елементу, причому завчасно момент відмови невідомо. Позначимо як T неперервну випадкову величину – час безвідмовної роботи елементу. Тоді подія $T<t$ полягає в тому, що за час t відбулася відмова елементу. А ймовірність цієї події є функцією розподілу НВВ T: $F(t)=P(T<t)$. Отже, ймовірність безвідмовної роботи елементу на протязі часу t: $P(T>t)=1-P(T<t)= \; =1-F(t)$.

Означення. *Функцією надійності* $R(t)$ називається функція, що визначає ймовірність безвідмовної роботи елементу на протязі часу t:

$$R(t)=P(T>t)=1-F(t).$$

Як правило, час безвідмовної роботи елементу має показниковий розподіл:

$$F(t)=1-e^{-\lambda t} \quad (\lambda>0;\; t\geq 0).$$

Тоді функція надійності:

$$R(t)=1-F(t)=1-(1-e^{-\lambda t})=e^{-\lambda t}.$$

Показниковим законом надійності називається функція надійності, яка визначається рівністю:

$$R(t)=e^{-\lambda t},$$

де $\lambda>0$ – інтенсивність відмов.

Приклад. Відбувається випробування 3-х елементів, що працюють незалежно один від одного. Час безвідмовної праці елементів розподілено за показниковим законом: для 1-го елемента щільність розподілу $f_1(t) = 0,1e^{-0,1t}$, для 2-го – $f_2(t) = 0,2e^{-0,2t}$, для 3-го – $f_3(t) = 0,3e^{-0,3t}$. Знайти ймовірність того, що в інтервалі часу $(0, 10)$ відмовить хоча б один елемент.

Знайдемо функцію розподілу для кожного елемента: $F_1(t) = 1 - e^{-0,1t}$; $F_2(t) = 1 - e^{-0,2t}$; $F_3(t) = 1 - e^{-0,3t}$.

Ймовірність того, що за інтервал часу $(0, 10)$ 1-й елемент буде працювати безвідмовно, дорівнює:

$$R_1(10) = 1 - F_1(10) = e^{-0,1\cdot 10} = e^{-1}.$$

Для 2-го та 3-го елементів відповідні ймовірності дорівнюють:

$$R_2(10) = 1 - F_2(10) = e^{-0,2\cdot 10} = e^{-2},$$

$$R_3(10) = 1 - F_3(10) = e^{-0,3\cdot 10} = e^{-3}.$$

Тоді ймовірність того, що за інтервал часу $(0, 10)$ відмовить хоча б один елемент, дорівнює:

$$1 - R_1(10)R_2(10)R_3(10) = 1 - e^{-1}e^{-2}e^{-3} = 1 - e^{-6} \approx 0,9975.$$

Показниковий закон надійності має наступну важливу властивість.

Теорема. *Ймовірність безвідмовної роботи елемента на інтервалі часу тривалістю t не залежить від часу попередньої роботи до початку інтервалу, що розглядається, а залежить тільки від тривалості часу t (при заданій інтенсивності відмов λ).*

Доведення. Нехай подія A – безвідмовна робота елемента на інтервалі $(0, t_0)$ тривалістю $(0, t_0)$. Подія B – безвідмовна робота елемента на інтервалі $(t_0, t_0 + t)$ тривалістю t. Тоді AB – безвідмовна робота елемента на інтервалі $(0, t_0 + t)$ тривалістю $t_0 + t$.

Знайдемо:

$$P(A) = e^{-\lambda t_0},\ \ P(B) = e^{-\lambda t}\ ,\quad P(AB) = e^{-\lambda(t_0+t)} = e^{-\lambda t_0}e^{-\lambda t}.$$

Знайдемо умовну ймовірність того, що елемент працюватиме безвідмовно на інтервалі $(t_0, t_0 + t)$ за умови, що він вже працював безвідмовно на попередньому інтервалі $(0, t_0)$:

$$P\left(B/A\right) = \frac{P(AB)}{P(A)} = \frac{e^{-\lambda t_0}e^{-\lambda t}}{e^{-\lambda t_0}} = e^{-\lambda t}.$$

Таким чином, ймовірність $P\left(B/A\right)$ не залежить від t_0, а залежить тільки від t. Це й означає, що час праці на попередньому інтервалі не впливає на ймовірність безвідмовної праці на наступному інтервалі, яка залежить тільки від тривалості t цього наступного інтервалу.

Зауваження. При доведенні цієї теореми ми суттєво скористалися тим, що закон надійності саме показниковий. Можна довести й обернене твердження, а саме, якщо ймовірність безвідмовної роботи елемента на деякому інтервалі часу не залежить від його роботи на попередніх інтервалах часу, то закон надійності має показниковий розподіл.

2.24. Нормальний розподіл

В практичних задачах досить часто доводиться мати справу з випадковими величинами, які мають тенденцію групування навколо деякого середнього значення. Наприклад, якщо замірювати зріст студентів курсу, можна помітити, що більшість буде мати зріст, близький до середнього. Хоча, напевно, будуть студенти, що мають зріст помітно більший середнього чи помітно менший, але таких студентів буде порівняно небагато. Такі випадкові величини називаються нормальними.

Означення. НВВ X називається розподіленою за *нормальним законом* з параметрами a і σ $(\sigma > 0)$, якщо щільність її розподілу має вигляд:

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

Перевіримо виконання умови нормування. Знайдемо:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \left[\begin{array}{c} \frac{x-a}{\sigma} = t, \;\; x = \sigma t + a, \;\; dx = \sigma dt \\ x = \pm\infty \Rightarrow t = \pm\infty \end{array} \right] =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{1}{\sigma} e^{-\frac{t^2}{2}} \sigma dt = \frac{2}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-\frac{t^2}{2}} dt = 1.$$

Тут скористалися інтегралом Пуассона (див. п. 1.16):

$$\frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{2}.$$

Знайдемо числові характеристики нормального розподілу. Почнемо з математичного сподівання:

$$M[X] = \int_{-\infty}^{+\infty} x f(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \frac{1}{\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \left[\begin{array}{c} \frac{x-a}{\sigma} = t, \;\; x = \sigma t + a, \;\; dx = \sigma dt \\ x = \pm\infty \Rightarrow t = \pm\infty \end{array} \right] =$$

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

Перший з інтегралів останньої суми є інтегралом у симетричних межах від непарної функції, отже він дорівнює нулю. Таким чином:

$$M[X]=\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\frac{2a}{\sqrt{2\pi}}\int_{0}^{+\infty}e^{-\frac{t^2}{2}}dt=a$$

(знову на підставі інтеграла Пуассона). Отже, *математичне сподівання нормального розподілу дорівнює параметру* a.

Знайдемо дисперсію. Цього разу користуємось загальною формулою для дисперсії, а не «спрощеною».

$$D[X]=\int_{-\infty}^{+\infty}(x-m)^2 f(x)dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2\frac{1}{\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$

$$\left[\begin{array}{c}\dfrac{x-a}{\sigma}=t,\ \ x=\sigma t+a,\ \ dx=\sigma dt\\ x=\pm\infty\Rightarrow t=\pm\infty\end{array}\right]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma^2t^2\frac{1}{\sigma}e^{-\frac{t^2}{2}}\sigma dt=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2e^{-\frac{t^2}{2}}dt=$$

$$=\frac{2\sigma^2}{\sqrt{2\pi}}\int_{0}^{+\infty}t^2e^{-\frac{t^2}{2}}dt=\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to+\infty}\int_{0}^{b}t^2e^{-\frac{t^2}{2}}dt=\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to+\infty}\int_{0}^{b}t\cdot te^{-\frac{t^2}{2}}dt=$$

$$=-\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to+\infty}\int_{0}^{b}td\left(e^{-\frac{t^2}{2}}\right)=-\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to+\infty}\left(te^{-\frac{t^2}{2}}\Big|_0^b-\int_{0}^{b}e^{-\frac{t^2}{2}}dt\right)=$$

$$=-\frac{2\sigma^2}{\sqrt{2\pi}}\lim_{b\to+\infty}\left(\frac{b}{e^{b^2/2}}-\int_{0}^{b}e^{-\frac{t^2}{2}}dt\right)=\frac{2\sigma^2}{\sqrt{2\pi}}\int_{0}^{+\infty}e^{-\frac{t^2}{2}}dt=\sigma^2.$$

Тут скористалися тим, що

$$\lim_{b\to+\infty}\frac{b}{e^{b^2/2}}=0,$$

та інтегралом Пуассона. Таким чином:

$$D[X]=\sigma^2,$$

тобто дисперсія нормального розподілу дорівнює σ^2. Отже, *середньоквадратичне відхилення нормального розподілу дорівнює параметру* σ:

$$\sigma[X]=\sqrt{D[X]}=\sqrt{\sigma^2}=\sigma.$$

Водночас ми встановили ймовірнісний зміст параметрів a і σ нормального розподілу. Параметр a – математичне сподівання цього розподілу, а параметр σ – середньоквадратичне відхилення.

Означення. Якщо НВВ X_0 розподілена нормально з параметрами $a=0$ і $\sigma=1$, то величина X_0 називається *нормованою*.

Щільність розподілу нормованої нормальної величини:

$$f_0(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Легко помітити, що $f_0(x)=\varphi(x)$ – диференціальна функція Лапласа (див. п. 1.15).

Функція розподілу нормованої величини має вигляд:

$$F_0(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$$

Її значення можна знайти, користуючись таблицями інтегральної функції Лапласа $\Phi(x)$ (див. п. 1.16). Дійсно:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{t^2}{2}}dt+\frac{1}{\sqrt{2\pi}}\int_{0}^{x}e^{-\frac{t^2}{2}}dt=0,5+\Phi(x).$$

Тут скористалися тим, що

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{t^2}{2}}dt\left[\begin{array}{c}\tau=-t,\ d\tau=-dt,\\ t=-\infty\Rightarrow\tau=+\infty,\ t=0\Rightarrow\tau=0\end{array}\right]=-\frac{1}{\sqrt{2\pi}}\int_{+\infty}^{0}e^{-\frac{\tau^2}{2}}d\tau=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{0}^{+\infty}e^{-\frac{\tau^2}{2}}d\tau=0,5$$

(інтеграл Пуассона).

Наприклад: $F_0(0,72)=0,5+\Phi(0,72)=0,5+0,2642=0,7642$.

Нехай X – нормальна випадкова величина з довільними параметрами a і σ. Розглянемо НВВ:

$$X_0=\frac{X-a}{\sigma}.$$

Ця величина також розподілена нормально. Покажемо, що вона є нормованою. Дійсно:

$$M[X_0]=M\left[\frac{X-a}{\sigma}\right]=\frac{1}{\sigma}M[X-a]=\frac{1}{\sigma}(M[X]-M[a])=\frac{1}{\sigma}(a-a)=0;$$

$$D[X_0]=D\left[\frac{X-a}{\sigma}\right]=\frac{1}{\sigma^2}D[X-a]=\frac{1}{\sigma^2}D[X]=\frac{1}{\sigma^2}\cdot\sigma^2=1,$$

отже, $\sigma[X_0]=1$.

Функція розподілу нормальної випадкової величини з довільними a і σ виражається через функцію $F_0(x)$ за формулою:

$$F(x)=F_0\left(\frac{x-a}{\sigma}\right).$$

Дійсно,

$$F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\frac{1}{\sigma}e^{-\frac{(t-a)^2}{2\sigma^2}}dt\left[\begin{array}{c}\frac{t-a}{\sigma}=\tau,\ t=\sigma\tau+a,\ dt=\sigma d\tau,\\ t=-\infty\Rightarrow\tau=-\infty,\ t=x\Rightarrow\tau=\frac{x-a}{\sigma}\end{array}\right]=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-a}{\sigma}}e^{-\frac{\tau^2}{2}}d\tau=F_0\left(\frac{x-a}{\sigma}\right).$$

Для відповідних щільностей розподілів маємо:

$$f(x)=\frac{1}{\sigma}f_0\left(\frac{x-a}{\sigma}\right).$$

2.25. Нормальна крива та вплив параметрів нормального розподілу на її форму

Означення. *Нормальною кривою* (або *кривою Гауса*) називається графік щільності нормального розподілу, тобто функції:

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

Для побудови цього графіку розглянемо спочатку випадок $a=0$, тобто функцію

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

Функція $z(x)$ визначена на всій числовій прямій. Ця функція приймає тільки додатні значення, і вона парна. Тому достатньо дослідити її поведінку лише при $x\ge 0$.

Маємо:

$$\lim_{x\to+\infty}z(x)=\lim_{x\to+\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}=0,$$

отже, вісь Ox є горизонтальною асимптотою графіка функції.

Дослідимо функцію на проміжки монотонності та точки екстремуму. Знайдемо:

$$z'(x)=\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}\right)'=\frac{1}{\sigma\sqrt{2\pi}}\left(-\frac{x}{\sigma^2}\right)e^{-\frac{x^2}{2\sigma^2}}=-\frac{x}{\sigma^3\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}.$$

Звідси випливає, що $z'(x) < 0$ при $x > 0$, тобто функція $z(x)$ при $x > 0$ є спадною. У точці $x = 0$ функція $z(x)$ має максимум, який дорівнює $1/(\sigma\sqrt{2\pi})$.

Дослідимо функцію на проміжки опуклості та вгнутості та точки перегину. Знайдемо:

$$z''(x) = \left(-\frac{x}{\sigma^3\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \right)' = -\frac{1}{\sigma^3\sqrt{2\pi}} \left(x e^{-\frac{x^2}{2\sigma^2}} \right)' =$$

$$= -\frac{1}{\sigma^3\sqrt{2\pi}} \left(e^{-\frac{x^2}{2\sigma^2}} - \frac{x^2}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \right) = \frac{1}{\sigma^3\sqrt{2\pi}} \left(\frac{x^2}{\sigma^2} - 1 \right) e^{-\frac{x^2}{2\sigma^2}}.$$

Звідси випливає, що на інтервалі $(0, +\infty)$ виконано: $z''(x) = 0$ при $x = \sigma$; $z''(x) < 0$ при $x \in (0, \sigma)$; $z''(x) > 0$ при $x \in (\sigma, +\infty)$. Отже, функція $z(x)$ є опуклою при $x \in (0, \sigma)$ і вгнутою при $x \in (\sigma, +\infty)$. Точка $x = \sigma$ є точкою перегину.

Таким чином, внаслідок парності графік функції $z(x)$ має вигляд:

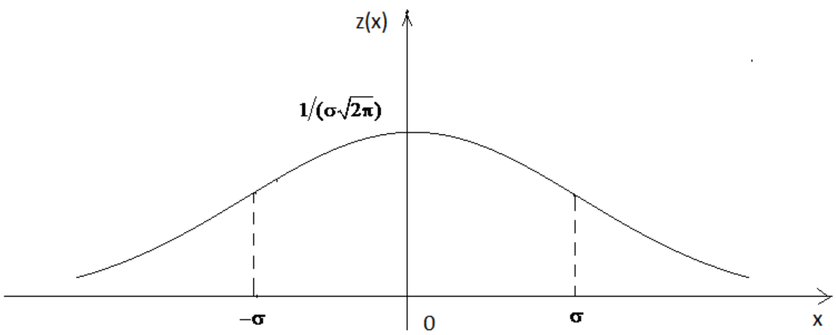


Рис. 26

Графік функції $y(x)$ можна отримати з графіка функції $z(x)$ паралельним зсувом вздовж осі Ox на a одиниць вправо, якщо $a > 0$, і на $-a$ одиниць вліво, якщо $a < 0$ (рис. 27). Цей графік й є нормальна крива або крива Гауса. Він симетричний відносно прямої $x = a$. Як ми знаємо, вже з цього випливає, що параметр a є математичним сподіванням нормального розподілу.

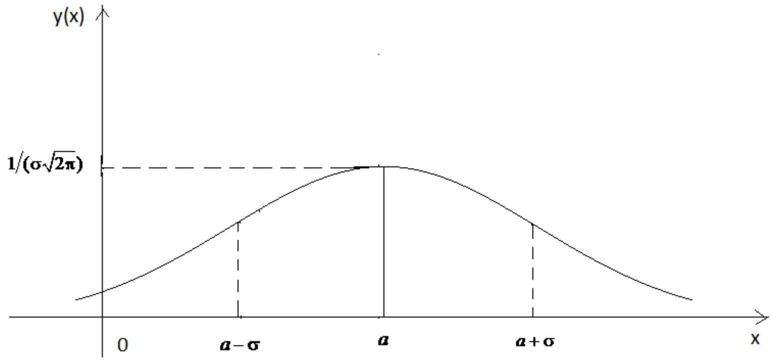


Рис. 27

З'ясуємо, як впливають на форму та розташування кривої Гауса зміна параметрів a і σ нормального розподілу. Зміна параметру a приводить лише до паралельного зсуву кривої вздовж осі Ox без зміни форми цієї кривої. А ось якщо змінюється параметр σ, то ситуація інша. Розглянемо границі

$$\lim_{\sigma\to 0}\frac{1}{\sigma\sqrt{2\pi}}=\infty,\quad \lim_{\sigma\to\infty}\frac{1}{\sigma\sqrt{2\pi}}=0.$$

Таким чином, при зменшенні параметра σ максимум $1/(\sigma\sqrt{2\pi})$ функції $y(x)$ збільшується, а при збільшенні σ цей максимум зменшується. Сама крива Гауса при зменшенні σ стає більш «гостровершинною», розтягується вздовж вертикальної прямої $x=a$. При збільшенні σ крива стає більш пологою, стискається до осі Ox (рис. 28).

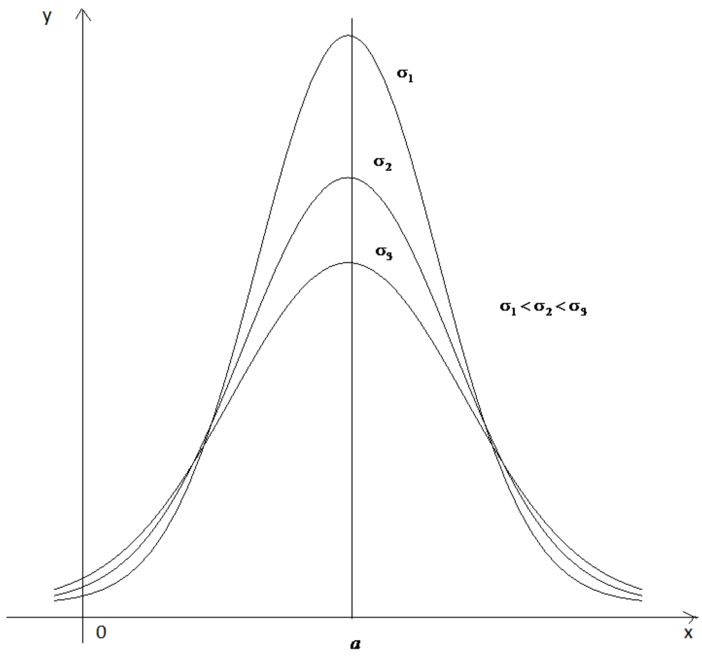


Рис. 28

2.26. Ймовірність потрапляння нормальної випадкової величини в заданий інтервал

Нехай НВВ X розподілена нормально з параметрами a і σ. Знайдемо $P(\alpha < X < \beta)$, де α, β – задані числа.

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)dx = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} \frac{1}{\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$$

$$\left[\begin{array}{l} \frac{x-a}{\sigma} = t,\ x = \sigma t + a,\ dx = \sigma dt, \\ x = \alpha \Rightarrow t = \frac{\alpha - a}{\sigma};\ x = \beta \Rightarrow t = \frac{\beta - a}{\sigma} \end{array} \right] = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} \frac{1}{\sigma} e^{-\frac{t^2}{2}} \sigma dt =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\alpha-a}{\sigma}} e^{-\frac{t^2}{2}} dt .$$

Користуючись інтегральною функцією Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt$$

(див. п. 1.16), остаточно отримуємо:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

Зокрема, якщо X_0 – нормована нормальна величина ($a = 0,\ \sigma = 1$), то:

$$P(\alpha < X_0 < \beta) = \Phi(\beta) - \Phi(\alpha).$$

Приклад. Математичне сподівання та середньоквадратичне відхилення нормально розподіленої випадкової величини X дорівнюють відповідно 10 і 2. Знайти ймовірність того, що внаслідок випробування величина X прийме значення з інтервалу (12, 14).

Маємо:

$$P(12 < X < 14) = \Phi\left(\frac{14-10}{2}\right) - \Phi\left(\frac{12-10}{2}\right) = \Phi(2) - \Phi(1) =$$

$$= 0{,}4772 - 0{,}3413 = 0{,}1359 .$$

2.27. Ймовірність заданого відхилення. Правило «трьох сигм»

Нехай НВВ X розподілена нормально з параметрами a і σ. Знайдемо ймовірність того, що внаслідок випробування X прийме значення, яке відрізняється від математичного сподівання не більше, ніж на задане число δ, тобто знайти ймовірність події $|X-a|<\delta$.

Скористаємось формулами попереднього параграфу:

$$P(|X-a|<\delta)=P(-\delta<X-a<\delta)=P(a-\delta<X<a+\delta)=$$

$$=\Phi\left(\frac{a+\delta-a}{\sigma}\right)-\Phi\left(\frac{a-\delta-a}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)-\Phi\left(-\frac{\delta}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)+\Phi\left(\frac{\delta}{\sigma}\right)=2\Phi\left(\frac{\delta}{\sigma}\right).$$

Тут скористалися властивістю непарності функції $\Phi(x)$: $\Phi(-x)=-\Phi(x)$. Отже, отримали формулу:

$$P(|X-a|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right). \qquad (*)$$

Зокрема, для нормованої величини X_0 ($a=0,\ \sigma=1$) дістанемо:

$$P(|X|<\delta)=2\Phi(\delta).$$

Покладемо у формулі (*) $\delta=\sigma t$. Дістанемо:

$$P(|X-a|<\sigma t)=2\Phi\left(\frac{\sigma t}{\sigma}\right)=2\Phi(t).$$

Якщо, зокрема, $t=3$, то

$$P(|X-a|<3\sigma)=2\Phi(3)\approx 2\cdot 0{,}49865=0{,}9973.$$

Подія, яка має таку ймовірність, є практично достовірною. Тобто можна практично напевно стверджувати, що внаслідок випробування нормальна величина X прийме значення з інтервалу $(a-3\sigma,\ a+3\sigma)$. Це твердження називається *правилом «трьох сигм».*

2.28. Задачі на нормальний розподіл

Задача 1. Систематична похибка утримання висоти літаком +20 м, а випадкова похибка розподілена нормально і має середньоквадратичне відхилення 75 м. Для польоту літаку виділено повітряний коридор висотою 100 м. Яка ймовірність того, що літак буде летіти нижче, всередині та вище цього коридору, якщо літаку задано висоту, що відповідає середині коридору?

Знайдемо ймовірність того, що літак буде летіти нижче коридору. Нехай X – випадкова величина, яка дорівнює похибці утримання висоти літаком. Тоді літак опиниться нижче коридору, якщо $X < -50$ (літаку задано висоту, яка відповідає середині коридору висотою 100 м). Отже:

$$P(X < -50) = P(-\infty < X < -50) = \Phi\left(\frac{-50-20}{75}\right) - \Phi(-\infty) = \Phi(+\infty) - \Phi\left(\frac{70}{75}\right) =$$

$$= 0,5 - \Phi(0,93) = 0,5 - 0,3238 = 0,1762.$$

Тут під символом $\Phi(\pm\infty)$ розуміємо $\lim\limits_{x\to\pm\infty} \Phi(x)$.

Ймовірність того, що літак буде всередині коридору, дорівнює:

$$P(-50 < X < 50) = \Phi\left(\frac{50-20}{75}\right) - \Phi\left(\frac{-50-20}{75}\right) = \Phi(0,4) + \Phi(0,93) =$$

$$= 0,1554 + 0,3238 = 0,4792.$$

Нарешті, ймовірність того, що літак буде летіти вище коридору:

$$P(50 < X < +\infty) = \Phi(+\infty) - \Phi\left(\frac{50-20}{75}\right) = \Phi(+\infty) - \Phi(0,4) =$$

$$= 0,5 - 0,1554 = 0,3446.$$

Задача 2. Математичне сподівання та середньоквадратичне відхилення нормально розподіленої випадкової величини X відповідно дорівнюють 20 і 5. Знайти ймовірність того, що внаслідок випробування випадкова величина X прийме значення з інтервалу (15, 25).

Згідно з формулою для ймовірності потрапляння нормальної величини в заданий інтервал маємо:

$$P(15 < X < 25) = \Phi\left(\frac{25-20}{5}\right) - \Phi\left(\frac{15-20}{5}\right) = \Phi(1) - \Phi(-1) = 2\Phi(1) =$$

$$= 2 \cdot 0,3413 = 0,6826.$$

Задача 3. Автомат штампує деталі. Контролюється довжина деталі X, яка розподілена нормально з математичним сподіванням (проектна довжина) 50 *мм*. Фактично довжина виготовлених деталей не менше 32 *мм* і не більше 68 *мм*. Знайти ймовірність того, що довжина навмання взятої деталі більше 55 *мм*.

Для розв'язання задачі необхідно знайти параметр σ. Оскільки задано межі для довжини деталі, то подія $32 < X < 68$ вважається достовірною. З іншого боку:

$$P(32 < X < 68) = \Phi\left(\frac{68-50}{\sigma}\right) - \Phi\left(\frac{32-50}{\sigma}\right) = \Phi\left(\frac{18}{\sigma}\right) - \Phi\left(-\frac{18}{\sigma}\right) = 2\Phi\left(\frac{18}{\sigma}\right) = 1,$$

отже:

$$\Phi\left(\frac{18}{\sigma}\right)=\frac{1}{2}.$$

Тому можна покласти $18/\sigma=5$, звідки $\sigma=3,6$. Тепер знайдемо:

$$P(55<X<68)=\Phi\left(\frac{68-50}{3,6}\right)-\Phi\left(\frac{55-50}{3,6}\right)=\Phi(5)-\Phi(1,39)=$$

$=0,5-0,4177=0,0823$.

Задача 4. Випадкові похибки вимірювань підпорядковано нормальному закону з середньоквадратичним відхиленням $\sigma=20$ *мм* і математичним сподіванням $a=0$. Знайти ймовірність того, що з 3-х незалежних випробувань похибка хоча б одного не перевищить за абсолютною величиною 4 *мм*.

Нехай НВВ X – випадкова помилка вимірювання. Знайдемо ймовірність p появи події $|X|<4$ у одному випробуванні:

$$p=P(|X|<4)=P(-4<X<4)=\Phi\left(\frac{4-0}{20}\right)-\Phi\left(\frac{-4-0}{20}\right)=2\Phi(0,2)=$$

$=2\cdot 0,0793=0,1586$.

Тоді:

$$1-p=P(|X|\geq 4)=1-0,1586=0,8414.$$

Тепер шукану ймовірність знайдемо за формулою:

$$P=1-(1-p)^3=1-0,8414^3\approx 0,404.$$

Задача 5. Деталь, що виготовлена автоматом, вважається придатною, якщо відхилення її розміру від проектного не перевищує 10 *мм*. Випадкові відхилення підпорядковані нормальному закону з середньоквадратичним відхиленням $\sigma=5$ *мм* і математичним сподіванням $a=0$. Скільки відсотків придатних деталей у середньому виготовлює автомат?

Нехай НВВ X – відхилення розміру деталі від проектного. За умовою ця величина розподілена нормально з параметрами $a=0$ і $\sigma=5$. Скористаємось формулою:

$$P(|X|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right).$$

Підставляючи $\delta=10$, $\sigma=5$, дістанемо:

$$P(|X|<10)=2\Phi\left(\frac{10}{5}\right)=2\Phi(2)=2\cdot 0,4772=0,9544.$$

Таким чином, ймовірність відхилення, меншого за 10 *мм,* дорівнює 0,9544. Звідси випливає, що приблизно 95% відсотків деталей придатні.

Задача 6. НВВ X розподілена нормально, причому $a=0$. Знайти значення σ, за яким ймовірність потрапляння значення X в інтервал (α,β), де $0<\alpha<\beta$, буде найбільшою.

Розглянемо функцію $\psi(\sigma)$, яка, з урахуванням того, що $a=0$, дорівнює:

$$\psi(\sigma)=P(\alpha<X<\beta)=\Phi\left(\frac{\beta}{\sigma}\right)-\Phi\left(\frac{\alpha}{\sigma}\right)=\frac{1}{\sqrt{2\pi}}\int_0^{\beta/\sigma}e^{-\frac{t^2}{2}}dt-\frac{1}{\sqrt{2\pi}}\int_0^{\alpha/\sigma}e^{-\frac{t^2}{2}}dt.$$

Знайдемо тепер σ з рівняння $\psi'(\sigma)=0$:

$$\psi'(\sigma)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(\beta/\sigma)^2}{2}}\left(-\frac{\beta}{\sigma^2}\right)-\frac{1}{\sqrt{2\pi}}e^{-\frac{(\alpha/\sigma)^2}{2}}\left(-\frac{\alpha}{\sigma^2}\right)=$$

$$=\frac{\alpha}{\sigma^2\sqrt{2\pi}}e^{-\frac{\alpha^2}{2\sigma^2}}-\frac{\beta}{\sigma^2\sqrt{2\pi}}e^{-\frac{\beta^2}{2\sigma^2}}=0.$$

Звідси:

$$\alpha e^{-\frac{\alpha^2}{2\sigma^2}}=\beta e^{-\frac{\beta^2}{2\sigma^2}},$$

або

$$e^{\frac{\beta^2-\alpha^2}{2\sigma^2}}=\frac{\beta}{\alpha}.$$

Звідси:

$$\frac{\beta^2-\alpha^2}{2\sigma^2}=\ln\frac{\beta}{\alpha}=\ln\beta-\ln\alpha,$$

і тоді

$$\sigma=\sqrt{\frac{\beta^2-\alpha^2}{2(\ln\beta-\ln\alpha)}}.$$

Далі треба переконатися в тому, що при знайденому значенні σ виконується нерівність $\psi''(\sigma)<0$, це й буде означати що при знайденому значенні σ досягається максимум функції $\psi(\sigma)$. Внаслідок громіздкості викладок ми це доведення опускаємо.

2.29. Відхилення заданого розподілу від нормального Асиметрія та ексцес

Випадкові величини, які реально зустрічаються на практиці, хоча можуть бути близькими до нормальних, але їх розподіл, взагалі кажучи, не збігається точно з нормальним, а дещо від нього відрізняється. І тому виникає необхідність ці відмінності оцінювати. З цією метою вводяться такі характеристики, як асиметрія та ексцес. Можна довести, що для нормального розподілу ці характеристики дорівнюють нулю. Тому якщо для реального розподілу вони є невеликими, то можна припустити, що даний розподіл близький до нормального.

Означення. *Асиметрією* даного розподілу називається відношення центрального моменту 3-го порядку даного розподілу до куба його середньоквадратичного відхилення:

$$A = \frac{\mu_3}{\sigma^3}.$$

Для нормального розподілу (і взагалі для будь-якого симетричного розподілу, зокрема, рівномірного) $A = 0$. Якщо $A \neq 0$, то графік даного розподілу вже не є симетричним відносно прямої $x = M[X]$. У випадку $A > 0$ частина графіка щільності, що розташована справа від прямої $x = M[X]$, є більш пологою, а частина, що розташована зліва, більш крутою (рис. 29).

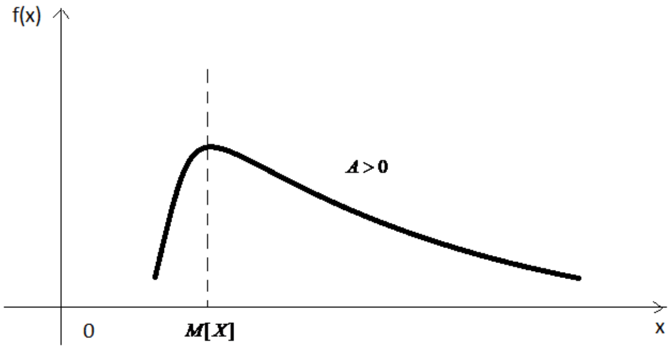


Рис. 29

У випадку $A < 0$ ситуація протилежна: ліва частина графіка більш полога, а права – більш крута (рис. 30).

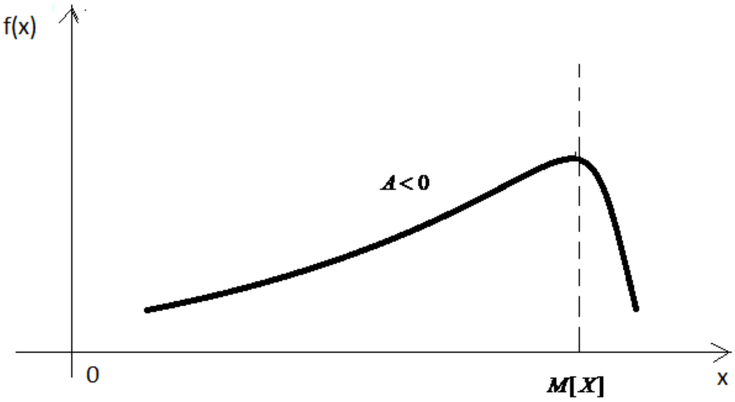


Рис. 30

Для оцінки «крутизни» графіка, тобто більшого чи меншого його підйому в околі математичного сподівання порівняно з нормальним розподілом, використовують таку характеристику, як ексцес.

Означення. *Ексцесом* даного розподілу називається характеристика, що визначається рівністю:

$$E = \frac{\mu_4}{\sigma^4} - 3.$$

Знову ж таки, для нормального розподілу $E = 0$. Тому, якщо $E \neq 0$, то даний розподіл відрізняється від нормального, але якщо $E \approx 0$, то можна вважати, що розподіл близький до нормального.

При $E > 0$ крива даного розподілу має більш високу і гостру верхівку, крива більш крута порівняно з нормальною. При $E < 0$ крива розподілу має більш низьку і пологу верхівку порівняно з нормальною (рис. 31).

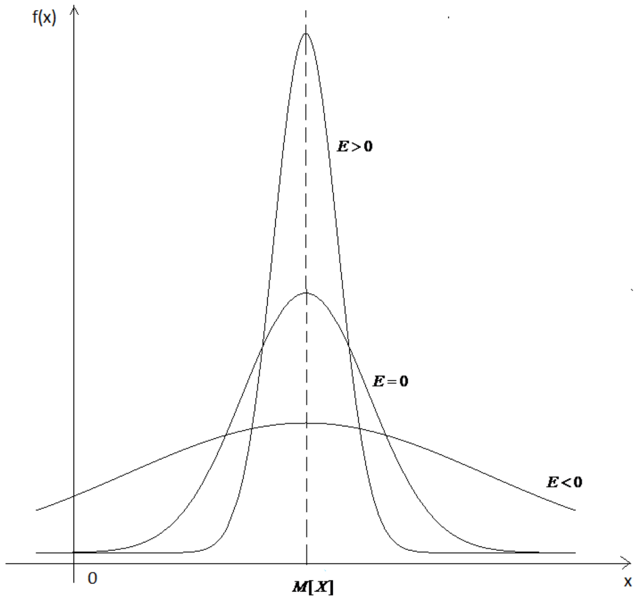


Рис. 31

2.30. Поняття про центральну граничну теорему Значення нормально розподілених випадкових величин

Нормально розподілені випадкові величини відіграють дуже помітну роль в різноманітних процесах. Нормальному закону підпорядковуються, наприклад, похибки вимірювань, оцінки потужності геологічних порід, вміст хімічних елементів у породах. Пояснюється це наступним фактом: *якщо ми маємо суму дуже великого числа взаємно незалежних випадкових величин, вплив кожної з яких на всю суму зовсім незначний, то вся сума має розподіл, близький до нормального.* Цей факт доведено строго математично і має назву *центральної граничної теореми* О. М. Ляпунова.

Нехай $X_1, X_2, \ldots, X_n, \ldots$ – послідовність взаємно незалежних випадкових величин, кожна з яких має скінченні математичне сподівання та дисперсію:

$$M[X_k] = a_k, \quad D[X_k] = b_k^2.$$

Позначимо:

$$S_n = \sum_{k=1}^{n} X_k, \quad A_n = \sum_{k=1}^{n} a_k, \quad B_n^2 = \sum_{k=1}^{n} b_k^2.$$

Очевидно тоді, що

$$M[S_n]=A_n,\quad D[S_n]=B_n^2,\quad \sigma[S_n]=B_n.$$

Випадкова величина

$$\frac{S_n-A_n}{B_n}$$

тоді є нормованою. Позначимо її функцію розподілу як $F_n(x)$:

$$F_n(x)=P\left(\frac{S_n-A_n}{B_n}<x\right).$$

Кажуть, що до послідовності $X_1,X_2,\ldots,X_n,\ldots$ можна застосувати центральну граничну теорему, якщо для будь-якого x виконано:

$$\lim_{n\to\infty}F_n(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$$

Тобто функція розподілу нормованої величини $(S_n-A_n)/B_n$ при $n\to\infty$ прямує до функції розподілу нормальної нормованої величини. Зокрема, якщо всі випадкові величини $X_1,X_2,\ldots,X_n,\ldots$ однаково розподілені, то до цієї послідовності можна застосувати центральну граничну теорему, якщо дисперсії всіх величин $X_k\ (k=1,2,\ldots)$ скінченні та відмінні від нуля. О. М. Ляпунов довів, що якщо відношення

$$L_n=\frac{C_n}{B_n^3},$$

де

$$C_n=\sum_{k=1}^{n}M\left[(X_k-a_k)^3\right],$$

прямує до нуля, то до послідовності $X_1,X_2,\ldots,X_n,\ldots$ можна застосувати центральну граничну теорему.

2.31. Деякі інші важливі розподіли

Розглянемо деякі інші розподіли, які зустрічаються в прикладних задачах.

1. Розподіл Лапласа

Цей розподіл визначається щільністю:

$$f(x)=\frac{1}{2}\lambda e^{-\lambda|x|},\quad -\infty<x<+\infty,$$

де a, λ $(\lambda>0)$ – параметри розподілу. Графік наведено на рис. 32.

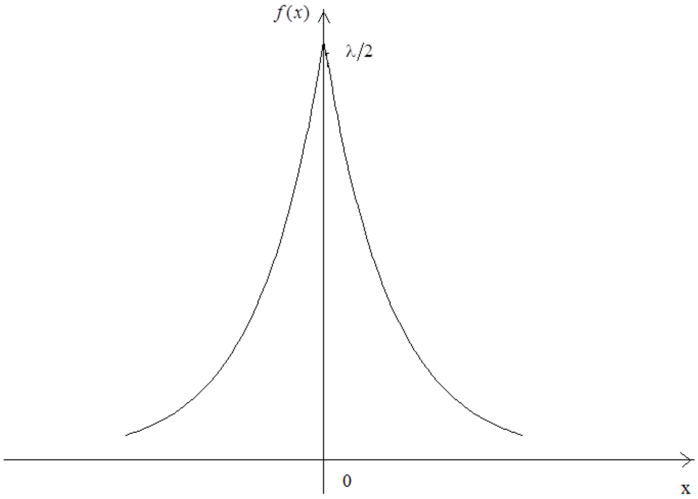


Рис. 32

Розподілу Лапласа підпорядковуються, наприклад, похибки в моделях регресії.

Числові характеристики розподілу Лапласа:

$M[X]=0$, $D[X]=\dfrac{2}{\lambda^2}$, $\sigma[X]=\dfrac{\sqrt{2}}{\lambda}$.

2. <u>Логнормальний розподіл</u>

Випадкова величина X називається *логарифмічно нормальною*, якщо $\ln X$ розподілено за нормальним законом. Логнормальний розподіл використовується, наприклад, при моделюванні таких змінних, як прибуток, допустиме відхилення від стандарту відсотків шкідливих речовин у харчових продуктах тощо. Щільність розподілу має вигляд:

$$f(x)=\frac{1}{\sigma x\sqrt{2\pi}}\exp\left(-\frac{(\ln x-\ln a)^2}{2\sigma^2}\right),\ x>0,$$

a, σ $(\sigma>0)$ – параметри. Графік наведено на рис. 33.

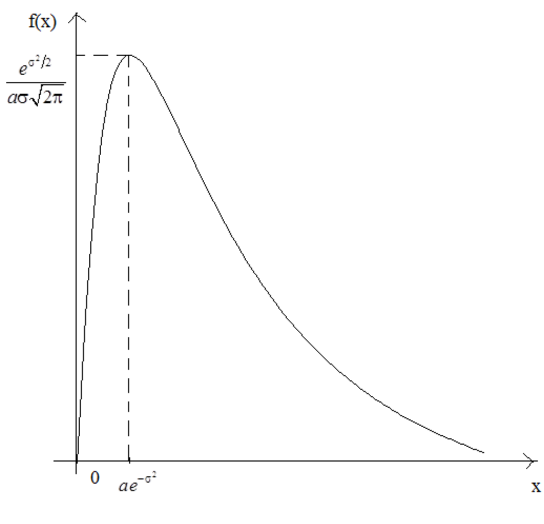


Рис. 33

Числові характеристики цього розподілу наступні:

$$M[X]=ae^{\sigma^2/2},\ \ D[X]=a^2e^{\sigma^2}\left(e^{\sigma^2}-1\right),\ \ \sigma[X]=ae^{\sigma^2/2}\sqrt{e^{\sigma^2}-1}\ .$$

3. <u>Розподіл арксинусу</u>

Випадкова величина X називається *розподіленою за законом арксинусу*, якщо її щільність розподілу має вигляд:

$$f(x)=\begin{cases}\dfrac{1}{\pi\sqrt{x(1-x)}},\ x\in(0,1),\\ 0,\ \ x\notin(0,1).\end{cases}$$

Графік щільності наведено на рис. 34.

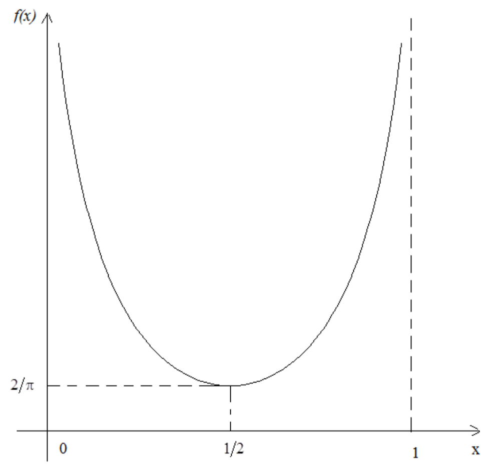


Рис. 34

Знайдемо функцію розподілу.

а) $x\le 0$: $F(x)=0$.

б) $0<x<1$:

$$F(x)=\int_0^x\frac{dt}{\pi\sqrt{t(1-t)}}=\frac{1}{\pi}\int_0^x\frac{dt}{\sqrt{t-t^2}}=\frac{1}{\pi}\int_0^x\frac{dt}{\sqrt{\dfrac{1}{4}-\left(t-\dfrac{1}{2}\right)^2}}=$$

$$=\frac{1}{\pi}\arcsin\frac{t-1/2}{1/2}\Bigg|_0^x=\frac{1}{\pi}\arcsin(2t-1)\Big|_0^x=\frac{1}{\pi}\arcsin(2x-1)+\frac{1}{2}.$$

в) $x\ge 1$: $F(x)=1$.

Отже

$$F(x)=\begin{cases}0,\ x\le 0,\\ \dfrac{1}{\pi}\arcsin(2x-1)+\dfrac{1}{2},\ 0<x<1,\\ 1,\ x\ge 1.\end{cases}$$

Графік функції $F(x)$ наведено на рис. 35.

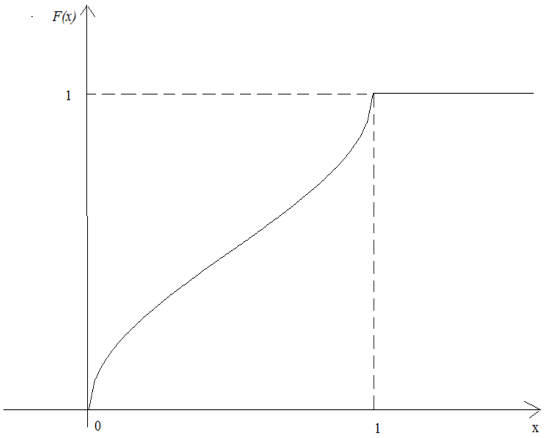


Рис. 35

Числові характеристики розподілу арксинусу:

$$M[X] = \frac{1}{2}, \quad D[X] = \frac{1}{8}, \quad \sigma[X] = \frac{1}{2\sqrt{2}}.$$

4. <u>Розподіл Коші</u>

Випадкова величина X називається *розподіленою за законом Коші*, якщо щільність розподілу має вигляд:

$$f(x) = \frac{a}{\pi\left(a^2 + x^2\right)}, \quad -\infty < x < +\infty.$$

Графік щільності наведено на рис. 36.

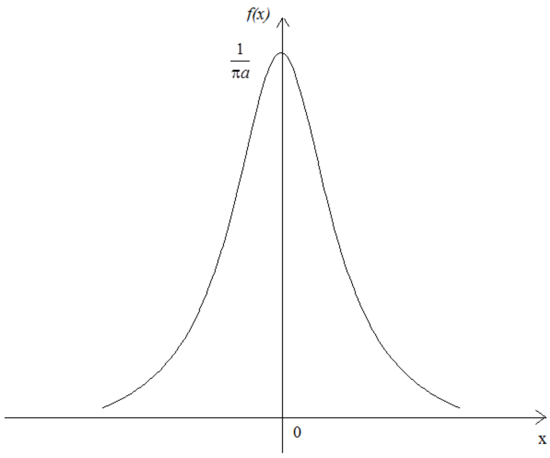


Рис. 36

Функція розподілу Коші така:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\frac{x}{a}, \quad -\infty < x < +\infty.$$

Графік наведено на рис. 37.

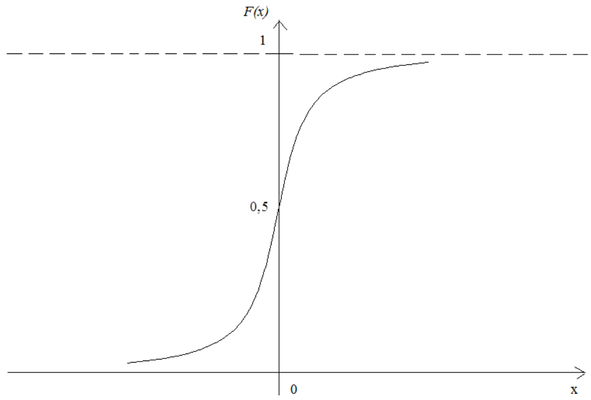


Рис. 37

А ось математичного сподівання та дисперсії розподілу Коші не існує внаслідок того, що інтеграли

$$\int_{-\infty}^{+\infty} \frac{|x|dx}{a^2+x^2}, \quad \int_{-\infty}^{+\infty} \frac{x^2dx}{a^2+x^2}$$

розбіжні.

2.32. Двовимірні випадкові величини
Закон розподілу дискретної двовимірної випадкової величини

У попередніх параграфах ми розглядали випадкові величини, можливі значення яких визначалися одним числом. Такі величини називаються *одновимірними*. Разом з цим зустрічаються випадкові величини, можливі значення яких визначаються декількома числами. Наприклад, з набору кісток доміно навмання вибирається одна кістка. На кожній кістці є два числа, наприклад, 2:5, або 6:0, або 3:3. Ці числа у сукупності можна розглядати як випадкову величину, можливі значення якої виражаються двома числами. Або нехай на координатну площину випадковим чином потрапляє точка. Положення цієї точки визначається двома незалежними числами – координатами цієї точки. Ці координати у сукупності також є випадковою величиною, можливі значення якої виражаються двома числами. Є й величини, можливі значення яких виражаються трьома, декількома числами. Такі величини називаються *багатовимірними.* Зокрема, *двовимірними, тривимірними*, тощо.

Будемо розглядати двовимірні випадкові величини. Позначатимемо такі величини як (X,Y). Кожну з величин X та Y називають *складовою* (або *компонентою*). Обидві величини X та Y, що розглядаються одночасно, утворюють систему двох випадкових величин. Її також можна інтерпретувати як вектор на площині з компонентами X та Y. Тому також використовується термін *випадковий вектор* $\{X, Y\}$.

Якщо величини X та Y дискретні, то двовимірна випадкова величина (X, Y) також називається *дискретною*.

Означення. *Законом розподілу* дискретної двовимірної випадкової величини (X, Y) називається таблиця, яка містить перелік всіх можливих значень (x_i, y_j) цієї величини, а також ймовірності p_{ij} цих значень.

Тобто: $p_{ij} = P(X = x_i,\ Y = y_j)$ $(i = 1,2,...,n;\ j = 1,2,...,m)$.

Y	X					
	x_1	x_2	$\cdots$	x_i	$\cdots$	x_n
y_1	p_{11}	p_{21}	$\cdots$	p_{i1}	$\cdots$	p_{n1}
y_2	p_{12}	p_{22}	$\cdots$	p_{i2}	$\cdots$	p_{n2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_j	p_{1j}	p_{2j}	$\cdots$	p_{ij}	$\cdots$	p_{nj}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m	p_{1m}	p_{2m}	$\cdots$	p_{im}	$\cdots$	p_{nm}

Події $(X = x_i,\ Y = y_j)$ $(i = 1,2,...,n;\ j = 1,2,...,m)$ утворюють повну групу. Тому

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} = 1$$

(умова нормування).

Знаючи закон розподілу двовимірної дискретної випадкової величини, можна знайти закони розподілу кожної з компонент. Дійсно, події, наприклад, $(X = x_1, Y = y_1)$, $(X = x_1, Y = y_2)$, …, $(X = x_1, Y = y_m)$ несумісні, тому ймовірність $P(X = x_1)$ за теоремою про ймовірність суми подій така:

$$P(X = x_1) = p_{11} + p_{12} + \ldots + p_{1m}.$$

Таким чином $P(X = x_1)$ дорівнює сумі ймовірностей, що містяться у першому стовпці таблиці розподілу. Відповідно $P(X = x_i)$ дорівнює сумі ймовірностей, що містяться у i-му стовпці таблиці. Аналогічно $P(Y = y_j)$ дорівнює сумі ймовірностей, що містяться у j-му рядку таблиці.

<u>Приклад</u>. Задано закон розподілу двовимірної випадкової величини:

Y	X			
	26	30	41	50
2,3	0,05	0,12	0,08	0,04
2,7	0,09	0,30	0,11	0,21

Знайти закони розподілу компонент X та Y.

Маємо:

$P(X = 26) = 0,05 + 0,09 = 0,14$;

$P(X = 30) = 0,12 + 0,30 = 0,42$;

$P(X = 41) = 0,08 + 0,11 = 0,19$;

$P(X = 50) = 0,04 + 0,21 = 0,25$;

$P(Y = 2,3) = 0,05 + 0,12 + 0,08 + 0,04 = 0,29$;

$P(Y = 2,7) = 0,09 + 0,30 + 0,11 + 0,21 = 0,71$.

Отже:

X	26	30	41	50
P	0,14	0,42	0,19	0,25

Y	2,3	2,7
P	0,29	0,71

2.33. Функція розподілу двовимірної випадкової величини

Розглянемо двовимірну випадкову величину (X, Y). Нехай x, y – пара дійсних чисел.

Означення. *Функцією розподілу* двовимірної випадкової величини (X, Y) називається функція $F(x, y)$, яка дорівнює ймовірності добутку подій $(X < x)$ та $(Y < y)$:

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

З геометричної точки зору $F(x, y)$ є ймовірність того, що випадкова точка потрапить у нескінченний прямокутник з вершиною у точці (x, y) та розміщений нижче та лівіше цієї вершини (прямі лінії, що утворюють верхню та праву межу цього прямокутника, до нього не включаються) (рис. 38).

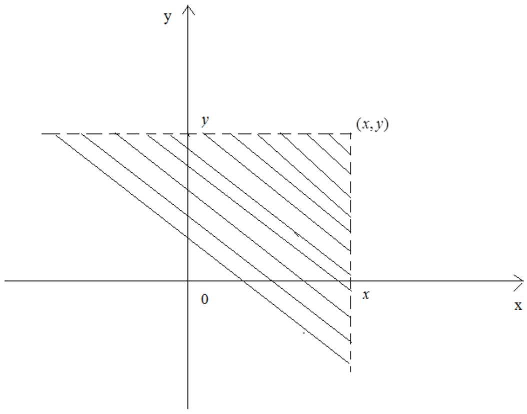


Рис. 38

<u>Приклад</u>. Знайти $P(X<2, Y<3)$, якщо

$$F(x,y)=\left(\frac{1}{\pi}\operatorname{arctg}\frac{x}{2}+\frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg}\frac{y}{3}+\frac{1}{2}\right).$$

Маємо:

$$P(X<2,Y<3)=F(2,3)=\left(\frac{1}{\pi}\operatorname{arctg}\frac{2}{2}+\frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg}\frac{3}{3}+\frac{1}{2}\right)=$$

$$=\left(\frac{1}{\pi}\cdot\frac{\pi}{4}+\frac{1}{2}\right)^2=\left(\frac{3}{4}\right)^2=\frac{9}{16}.$$

З'ясуємо властивості функції розподілу двовимірної випадкової величини. Вони певною мірою аналогічні властивостям функції розподілу одновимірної випадкової величини.

Властивість 1. *Всі значення функції розподілу задовольняють нерівність:*

$0 \le F(x,y) \le 1$.

Доведення випливає з того, що функція розподілу за означенням є ймовірність, а будь-яка ймовірність є число з відрізку $[0,1]$.

Властивість 2. *Функція* $F(x,y)$ *неспадна за кожним аргументом, тобто*

$F(x_2,y) \ge F(x_1,y)$, *якщо* $x_2 > x_1$;

$F(x,y_2) \ge F(x,y_1)$, *якщо* $y_2 > y_1$.

Доведення. Доведемо, що функція $F(x,y)$ неспадна за аргументом x. Розглянемо подію:

$$(X<x_2,Y<y)=(X<x_1,Y<y)+(x_1\le X<x_2,Y<y).$$

За теоремою про ймовірність суми несумісних подій:

$$P(X<x_2,Y<y)=P(X<x_1,Y<y)+P(x_1\le X<x_2,Y<y)\ge P(X<x_1,Y<y),$$

оскільки $P(x_1 \le X < x_2, Y < y) \ge 0$. Таким чином:

$$F(x_2,y)\ge F(x_1,y),$$

що й треба було довести.

Аналогічно доводиться, що функція $F(x, y)$ неспадна також за аргументом y.

Властивість 3. *Справедливі граничні співвідношення:*

1) $F(-\infty, y) = 0$; 2) $F(x, -\infty) = 0$; 3) $F(-\infty, -\infty) = 0$; 4) $F(+\infty, +\infty) = 1$.

Доведення. 1) $F(-\infty, y)$ є ймовірність події $(X < -\infty, Y < y)$, але така подія неможлива, оскільки неможлива подія $(X < -\infty)$, отже, ймовірність цієї події дорівнює нулю.

2) $F(x, -\infty) = P(X < x, Y < -\infty) = 0$, оскільки подія $(Y < -\infty)$ неможлива.

3) $P(X < -\infty,\ Y < -\infty) = 0$, оскільки подія $(X < -\infty,\ Y < -\infty)$ неможлива.

4) $F(+\infty, +\infty) = P(X < +\infty,\ Y < +\infty) = 1$, оскільки події $(X < +\infty)$, $(Y < +\infty)$ достовірні.

Властивість 4. *При $y = +\infty$ функція розподілу двовимірної випадкової величини (X, Y) перетворюється на функцію розподілу складової X:*

$F(x, +\infty) = F_1(x)$.

При $x = +\infty$ функція розподілу двовимірної випадкової величини (X, Y) перетворюється на функцію розподілу складової Y:

$F(+\infty, y) = F_2(y)$.

Доведення. Розглянемо:

$F(x, +\infty) = P(X < x, Y < +\infty) = P(X < x)P(Y < +\infty) = F_1(x) \cdot 1 = F_1(x)$;

$F(+\infty, y) = P(X < +\infty, Y < y) = P(X < +\infty)P(Y < y) = 1 \cdot F_2(y) = F_2(y)$.

2.34. Щільність розподілу неперервної двовимірної випадкової величини

Означення. *Щільністю розподілу* неперервної двовимірної випадкової величини (X, Y) називається невід'ємна функція:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y},$$

де $F(x, y)$ – функція розподілу величини (X, Y).

<u>Приклад</u>. Функцію розподілу двовимірної випадкової величини (X, Y) задано формулою:

$$F(x, y) = \begin{cases} 1 - 3^{-x} - 3^{-y} + 3^{-x-y}, \ \textit{при}\ x \ge 0,\ y \ge 0, \\ 0, \ \textit{при}\ x < 0, \ \textit{або}\ \ y < 0. \end{cases}$$

Знайти щільність $f(x, y)$ даного розподілу.

Знайдемо частинні похідні від $F(x, y)$. При $x \ge 0,\ y \ge 0$ маємо:

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

Таким чином:

$$f(x,y)=\begin{cases}3^{-x-y}\ln^2 3,\ \textit{при}\ x\ge 0,\ y\ge 0,\\ \ 0,\ \textit{при}\ x<0,\ \textit{або}\ \ y<0.\end{cases}$$

Знаючи щільність розподілу, можна знайти функцію розподілу за формулою:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(t,\tau)dt\,d\tau\,.$$

<u>Приклад.</u> Нехай щільність розподілу задано формулою:

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

Знайти функцію розподілу:

Маємо:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}\frac{dt\,d\tau}{\pi^2(1+t^2)(1+\tau^2)}=\frac{1}{\pi^2}\int_{-\infty}^{x}\left(\frac{1}{1+\tau^2}\int_{-\infty}^{y}\frac{dt}{1+t^2}\right)d\tau=$$

$$=\frac{1}{\pi^2}\int_{-\infty}^{x}\frac{1}{1+\tau^2}\left(\operatorname{arctg}y+\frac{\pi}{2}\right)d\tau=\left(\frac{1}{\pi}\operatorname{arctg}y+\frac{1}{2}\right)\frac{1}{\pi}\int_{-\infty}^{x}\frac{d\tau}{1+\tau^2}=$$

$$=\left(\frac{1}{\pi}\operatorname{arctg}y+\frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg}x+\frac{1}{2}\right).$$

Нехай задано довільну область D на площині xOy. Тоді ймовірність потрапляння випадкової точки (X,Y) в область D визначається формулою:

$$P((X,Y)\in D)=\iint_D f(x,y)\,dx\,dy\,.$$

Звідси, зокрема, буде випливати умова нормування для щільності розподілу:

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

Зокрема, якщо $f(x,y)$ відмінна від нуля лише в деякій області D_1, то

$$\iint_{D_1}f(x,y)\,dx\,dy=1.$$

<u>Приклад</u>. Щільність розподілу двовимірної випадкової величини (X,Y) задається так: $f(x,y)=C\left(2-\sqrt{x^2+y^2}\right)$ в крузі $x^2+y^2\le 4$. А зовні цього круга $f(x,y)=0$. Знайти сталу C та ймовірність потрапляння випадкової точки (X,Y) в круг $x^2+y^2\le 1$.

Скористаємось умовою нормування. Маємо:

$$\iint_{D_1} C\left(2-\sqrt{x^2+y^2}\right)dxdy=1,$$

де

$$D_1=\left\{(x,y):x^2+y^2\le 4\right\}.$$

Або

$$C\iint_{D_1}\left(2-\sqrt{x^2+y^2}\right)dxdy=1.$$

Обчислимо інтеграл:

$$I=\iint_{D_1}\left(2-\sqrt{x^2+y^2}\right)dxdy=\iint_{x^2+y^2\le 4}\left(2-\sqrt{x^2+y^2}\right)dxdy.$$

Перейдемо у цьому інтегралі до полярних координат:

$$x=r\cos\varphi,\ y=r\sin\varphi,\ 0\le\varphi\le 2\pi,\ 0\le r\le 2.$$

Дістанемо:

$$I=\int_0^{2\pi}d\varphi\int_0^2(2-r)r\,dr=2\pi\int_0^2(2r-r^2)dr=2\pi\left(r^2\Big|_0^2-\frac{r^3}{3}\Big|_0^2\right)=2\pi\left(4-\frac{8}{3}\right)=\frac{8\pi}{3}.$$

Звідси:

$$C=\frac{1}{8\pi/3}=\frac{3}{8\pi}.$$

Таким чином:

$$f(x,y)=\frac{3}{8\pi}\left(2-\sqrt{x^2+y^2}\right)$$

в крузі $x^2+y^2\le 4$. Знайдемо тепер ймовірність потрапляння точки (X,Y) в круг $D_2=\left\{(x,y):x^2+y^2\le 1\right\}$:

$$P((X,Y)\in D_2)=\iint_{x^2+y^2\le 1}\frac{3}{8\pi}\left(2-\sqrt{x^2+y^2}\right)dxdy$$

$$\begin{bmatrix} x=r\cos\varphi,\ y=r\sin\varphi, \\ 0\le\varphi\le 2\pi,\ 0\le r\le 1 \end{bmatrix}==\frac{3}{8\pi}\int_0^{2\pi}d\varphi\int_0^1(2-r)r\,dr=\frac{3}{8\pi}2\pi\int_0^1(2r-r^2)dr=$$

$$=\frac{3}{4}\left(r^2\Big|_0^1-\frac{r^3}{3}\Big|_0^1\right)=\frac{3}{4}\left(1-\frac{1}{3}\right)=\frac{1}{2}.$$

Якщо відомо щільність $f(x,y)$ сумісного розподілу системи (X,Y), то можна знайти щільності розподілу компонент X та Y. Нехай $F_1(x)$ – функція розподілу компоненти X. Тоді щільність $f_1(x)$ розподілу компоненти X дорівнює:

$$f_1(x) = \frac{dF_1(x)}{dx} .$$

З огляду на співвідношення

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dxdy, \quad F_1(x) = F(x, +\infty),$$

знайдемо:

$$F_1(x) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f(x, y)\, dxdy .$$

Звідси:

$$\frac{dF_1(x)}{dx} = \int_{-\infty}^{+\infty} f(x, y)\, dy ,$$

або

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy .$$

Аналогічно:

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y)\, dx .$$

Приклад. Двовимірну випадкову величину (X, Y) задано щільністю розподілу:

$$f(x, y) = \begin{cases} 1/6\pi, \ \textit{якщо} \ x^2/9 + y^2/4 \le 1, \\ 0, \ \textit{якщо} \ x^2/9 + y^2/4 > 1. \end{cases}$$

Знайти щільності розподілу складових X і Y.

Знайдемо:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy = \frac{1}{6\pi} \int_{-2\sqrt{1-x^2/9}}^{2\sqrt{1-x^2/9}} dy = \frac{1}{3\pi} \int_{0}^{2\sqrt{1-x^2/9}} dy = \frac{2}{3\pi} \sqrt{1 - \frac{x^2}{9}} = \frac{2}{9\pi} \sqrt{9 - x^2} .$$

Отже:

$$f_1(x) = \begin{cases} 2\sqrt{9 - x^2} / 9\pi, \ \textit{якщо} \ |x| \le 3, \\ 0, \ \textit{якщо} \ |x| > 3. \end{cases}$$

Аналогічно

$$f_2(y) = \begin{cases} \sqrt{4 - y^2} / 2\pi, \ \textit{якщо} \ |y| \le 2, \\ 0, \ \textit{якщо} \ |y| > 2. \end{cases}$$

2.35. Залежні та незалежні випадкові величини

Ми вже знаємо (див. п. 2.5), що дві випадкові величини X і Y називаються незалежними, якщо закон розподілу однієї з них не залежить від того, які можливі значення прийняла інша величина. Знайдемо необхідні та достатні умови незалежності випадкових величин.

Теорема. *Для того, щоб випадкові величини X і Y були незалежними, необхідно і достатньо, щоб функція розподілу системи (X,Y) дорівнювала добутку функцій розподілу складових:*

$$F(x,y) = F_1(x)F_2(y).$$

Доведення. Необхідність. Нехай X і Y незалежні. Тоді події $X < x$ та $Y < y$ також незалежні, отже

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

або

$$F(x,y) = F_1(x)F_2(y),$$

і, таким чином, необхідність доведено.

Достатність. Нехай $F(x,y) = F_1(x)F_2(y)$. Звідси:

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

Тобто ймовірність сумісної появи подій $X < x$ та $Y < y$ дорівнює добутку ймовірностей цих подій. Це означає, що ці події незалежні, отже, незалежні величини X *і* Y.

Наслідок. *Для того, щоб неперервні випадкові величини X і Y були незалежними, необхідно і достатньо, щоб щільність розподілу системи (X,Y) дорівнювала добутку щільностей розподілів складових:*

$$f(x,y) = f_1(x)f_2(y).$$

Доведення. Необхідність. Нехай X і Y – незалежні неперервні випадкові величини. Тоді на підставі попередньої теореми:

$$F(x,y) = F_1(x)F_2(y).$$

Диференцюючи цю рівність спочатку за x, а потім за y, дістанемо:

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{dF_1(x)}{dx} \cdot \frac{dF_2(y)}{dy}.$$

Або:

$$f(x,y) = f_1(x)f_2(y),$$

і необхідність доведено.

Достатність. Нехай

$$f(x,y) = f_1(x)f_2(y).$$

Інтегруючи цю рівність за x і за y, дістанемо:

$$\int_{-\infty}^{y}\int_{-\infty}^{x} f(x,y)\,dxdy = \int_{-\infty}^{x} f_1(x)\,dx \int_{-\infty}^{y} f_2(y)\,dy,$$

або

$$F(x,y) = F_1(x)F_2(y).$$

Звідси на підставі попередньої теореми випливає, що величини X та Y незалежні.

2.36. Числові характеристики системи двох випадкових величин

Для опису системи двох випадкових величин використовують також інтегральні її характеристики. До них відносяться математичні сподівання та дисперсії компонент, а також інші характеристики – кореляційний момент і коефіцієнт кореляції.

Означення. *Кореляційним моментом* μ_{xy} системи двох випадкових величин (X,Y) називається математичне сподівання добутку відхилень цих величин від своїх математичних сподівань:

$$\mu_{xy} = M[(X - M[X])(Y - M[Y])].$$

Зокрема, якщо величини X, Y є дискретними, то

$$\mu_{xy} = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - M[X])(y_j - M[Y])p_{ij}.$$

А якщо величини X, Y є неперервними, то

$$\mu_{xy} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - M[X])(y - M[Y])f(x,y)\,dxdy.$$

Кореляційний момент характеризує зв’язок між величинами X та Y.

Зауваження. Кореляційний момент може бути записано у вигляді:

$$\mu_{xy} = M[XY] - M[X]M[Y].$$

Дійсно:

$$\mu_{xy} = M[XY - M[X]Y - M[Y]X + M[X]M[Y]] =$$

$$= M[XY] - M[X]M[Y] - M[Y]M[X] + M[M[X]M[Y]] =$$

$$= M[XY] - 2M[X]M[Y] + M[X]M[Y] = M[XY] - M[X]M[Y].$$

Теорема. *Кореляційний момент двох незалежних випадкових величин X та Y дорівнює нулю.*

Доведення. Оскільки X та Y незалежні, то

$$M[XY] = M[X]M[Y].$$

А тоді

$$\mu_{xy} = M[XY] - M[X]M[Y] = M[X]M[Y] - M[X]M[Y] = 0.$$

Кореляційний момент має розмірність, що дорівнює добутку розмірностей величин X та Y. Отже, величина кореляційного моменту залежить від одиниць виміру випадкових величин. За цією причиною для одних й тих ж випадкових величин величина кореляційного моменту може мати різні значення в залежності від того, в яких одиницях виміряні величини X та Y. Така особливість кореляційного моменту створює незручності при користуванні ним, отже, є його недоліком. Для подолання цього недоліку вводиться інша числова характеристика – коефіцієнт кореляції.

Означення. *Коефіцієнтом кореляції* r_{xy} випадкових величин X та Y називається відношення кореляційного моменту цих величин до добутку їх середньоквадратичних відхилень:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y}.$$

Розмірність μ_{xy} дорівнює добутку розмірностей величин X та Y, σ_x має розмірність величини X, а σ_y має розмірність величини Y. Отже r_{xy} – безрозмірна величина.

Якщо X та Y незалежні, то $r_{xy} = 0$, оскільки $\mu_{xy} = 0$.

Теорема. *Абсолютна величина кореляційного моменту двох випадкових величин X та Y не перевищує добутку їх середньоквадратичних відхилень:*

$$|\mu_{xy}| \le \sigma_x \sigma_y. \qquad (*)$$

Доведення. Введемо до розгляду випадкову величину $Z_1 = \sigma_y X - \sigma_x Y$ і знайдемо її дисперсію:

$$D[Z_1] = M\left[(Z_1 - M[Z_1])^2\right] = M\left[Z_1^2\right] - M^2[Z_1] = M\left[(\sigma_y X - \sigma_x Y)^2\right] - M^2[\sigma_y X - \sigma_x Y] =$$

$$= M\left[\sigma_y^2 X^2 - 2\sigma_x\sigma_y XY + \sigma_x^2 Y^2\right] - (\sigma_y M[X] - \sigma_x M[Y])^2 =$$

$$= \sigma_y^2 M\left[X^2\right] - 2\sigma_x\sigma_y M[XY] + \sigma_x^2 M\left[Y^2\right] - \sigma_y^2 M^2[X] + 2\sigma_x\sigma_y M[X]M[Y] - \sigma_x^2 M^2[Y] =$$

$$= \sigma_y^2\left(M\left[X^2\right] - M^2[X]\right) + \sigma_x^2\left(M\left[Y^2\right] - M^2[Y]\right) - 2\sigma_x\sigma_y\left(M[XY] - M[X]M[Y]\right) =$$

$$= \sigma_y^2 D[X] + \sigma_x^2 D[Y] - 2\sigma_x\sigma_y\mu_{xy} = \sigma_y^2\sigma_x^2 + \sigma_x^2\sigma_y^2 - 2\sigma_x\sigma_y\mu_{xy} = 2\sigma_x^2\sigma_y^2 - 2\sigma_x\sigma_y\mu_{xy}.$$

Оскільки будь-яка дисперсія невід’ємна, то

$$2\sigma_x^2\sigma_y^2 - 2\sigma_x\sigma_y\mu_{xy} \ge 0,$$

звідки

$$\mu_{xy} \le \sigma_x\sigma_y.$$

Вводячи випадкову величину $Z_2 = \sigma_y X + \sigma_x Y$, аналогічно знайдемо:

$\mu_{xy} \geq -\sigma_x \sigma_y$.

Таким чином, μ_{xy} задовольняє подвійну нерівність:

$-\sigma_x \sigma_y \leq \mu_{xy} \leq \sigma_x \sigma_y$,

яка еквівалентна нерівності:

$|\mu_{xy}| \leq \sigma_x \sigma_y$,

що й треба було довести.

Наслідок. *Абсолютна величина коефіцієнту кореляції не перевищує одиниці:*

$|r_{xy}| \leq 1$.

Для доведення достатньо поділити обидві частини нерівності (*) на $\sigma_x \sigma_y$.

Як і кореляційний момент, коефіцієнт кореляції випадкових величин X та Y характеризує зв'язок між цими величинами. Чим модуль коефіцієнта кореляції ближче до одиниці, тим цей зв'язок більш суттєвий.

Приклад. Щільність розподілу двовимірної випадкової величини (X, Y) задано формулою:

$$f(x,y) = \begin{cases} x + y, \ (x,y) \in D, \\ 0, \ (x,y) \notin D, \end{cases}$$

де

$$D = \{(x,y) : 0 \leq x \leq 1, \ 0 \leq y \leq 1\}.$$

Знайти $f_1(x), f_2(y), M[X], M[Y], \sigma_x, \sigma_y, \mu_{xy}, r_{xy}$.

Щільності компонент знаходимо за формулами:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy, \ f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx .$$

В даному випадку маємо:

$$f_1(x) = \int_0^1 (x+y)\,dy = \left.\frac{(x+y)^2}{2}\right|_{y=0}^{y=1} = \frac{(x+1)^2}{2} - \frac{x^2}{2} = x + \frac{1}{2}, \quad 0 \leq x \leq 1,$$

тобто

$$f_1(x) = \begin{cases} x + 1/2, \ 0 \leq x \leq 1, \\ 0, \ x < 0, \ x > 1. \end{cases}$$

Аналогічно:

$$f_2(y) = \begin{cases} y + 1/2, \ 0 \leq y \leq 1, \\ 0, \ y < 0, \ y > 1. \end{cases}$$

Знайдемо математичні сподівання компонент:

$$M[X]=\int_{-\infty}^{+\infty} x f_1(x)\,dx=\int_0^1 x\left(x+\frac{1}{2}\right)dx=\int_0^1\left(x^2+\frac{x}{2}\right)dx=\left.\frac{x^3}{3}\right|_0^1+\left.\frac{x^2}{4}\right|_0^1=\frac{7}{12}.$$

Аналогічно:

$$M[Y]=\frac{7}{12}.$$

Знайдемо дисперсії:

$$D[X]=\int_{-\infty}^{+\infty} x^2 f_1(x)\,dx-M^2[X]=\int_0^1 x^2\left(x+\frac{1}{2}\right)dx-\frac{49}{144}=\int_0^1\left(x^3+\frac{x^2}{2}\right)dx-\frac{49}{144}=$$

$$=\left.\frac{x^4}{4}\right|_0^1+\left.\frac{x^3}{6}\right|_0^1-\frac{49}{144}=\frac{1}{4}+\frac{1}{6}-\frac{49}{144}=\frac{1}{144}\ .$$

Отже:

$$\sigma[X]=\sigma_x=\frac{1}{12}\ .$$

Аналогічно:

$$\sigma[Y]=\sigma_y=\frac{1}{12}.$$

Знайдемо кореляційний момент:

$$\mu_{xy}=M[XY]-M[X]M[Y]=\iint_D xy(x+y)\,dx\,dy-\frac{49}{144}=$$

$$=\int_0^1 dx\int_0^1 xy(x+y)-\frac{49}{144}=\int_0^1 xdx\int_0^1 y(x+y)\,dy-\frac{49}{144}=\int_0^1 xdx\int_0^1\left(xy+y^2\right)dy-\frac{49}{144}=$$

$$=\int_0^1 x\left(x\left.\frac{y^2}{2}\right|_{y=0}^{y=1}+\left.\frac{y^3}{3}\right|_{y=0}^{y=1}\right)dx-\frac{49}{144}=\int_0^1 x\left(\frac{x}{2}+\frac{1}{3}\right)dx-\frac{49}{144}=\left.\frac{x^3}{6}\right|_0^1+\left.\frac{x^2}{6}\right|_0^1-\frac{49}{144}=-\frac{1}{144}.$$

Тепер знайдемо коефіцієнт кореляції:

$$r_{xy}=\frac{\mu_{xy}}{\sigma_x\sigma_y}=-\frac{1/144}{1/12\cdot 1/12}=-1.$$

2.37. Корельованість та залежність випадкових величин

Означення. Дві випадкові величини X та Y називаються *корельованими*, якщо їх кореляційний момент відмінний від нуля (це еквівалентно тому, що коефіцієнт кореляції відмінний від нуля). Відповідно, величини X та Y називаються *некорельованими*, якщо їх кореляційний момент дорівнює нулю.

Дві корельовані величини також є й залежними. Дійсно, якби вони були незалежні, то їх кореляційний момент дорівнював би нулю (див. п. 2.36), а це суперечить тому, що величини корельовано.

Чи можна стверджувати обернене, тобто що з залежності величин випливає їх корельованість? З'ясовується, що не завжди. Кореляційний момент залежних величин може бути відмінним від нуля, але може й дорівнювати нулю.

Приклад. Нехай двовимірну випадкову величину (X,Y) задано щільністю розподілу (див. п. 2.34):

$$f(x,y)=\begin{cases} 1/6\pi, \text{ якщо } x^2/9+y^2/4\le 1, \\ 0, \text{ якщо } x^2/9+y^2/4>1. \end{cases}$$

Доведемо, що величини X та Y залежні. В п. 2.34 було обчислено щільності розподілу компонент X та Y:

$$f_1(x)=\begin{cases} 2\sqrt{9-x^2}/9\pi, \text{ якщо } |x|\le 3, \\ 0, \text{ якщо } |x|>3. \end{cases}$$

$$f_2(y)=\begin{cases} \sqrt{4-y^2}/2\pi, \text{ якщо } |y|\le 2, \\ 0, \text{ якщо } |y|>2. \end{cases}$$

Оскільки $f(x,y)\ne f_1(x)f_2(y)$, то величини X та Y залежні (див. п. 2.35). Знайдемо кореляційний момент цих величин за формулою:

$$\mu_{xy}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-M[X])(y-M[Y])f(x,y)\,dxdy.$$

Внаслідок парності функцій $f_1(x)$, $f_2(y)$, отже симетрії їх графіків відносно прямої $x=0$ для функції $f_1(x)$ і прямої $y=0$ для функції $f_2(y)$, можемо зробити висновок, що $M[X]=M[Y]=0$, отже:

$$\mu_{xy}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\,dx\,dy.$$

Оскільки $f(x,y)$ відмінна від нуля тільки всередині еліпсу $x^2/9+y^2/4\le 1$, то

$$\mu_{xy} = \iint\limits_{x^2/9+y^2/4\le 1} xyf(x,y)\,dx\,dy = \frac{1}{6\pi}\iint\limits_{x^2/9+y^2/4\le 1} xy\,dx\,dy = \int\limits_{-3}^{3} xdx \int\limits_{-2\sqrt{1-x^2/9}}^{2\sqrt{1-x^2/9}} ydy\,.$$

Інтеграл

$$\int\limits_{-2\sqrt{1-x^2/9}}^{2\sqrt{1-x^2/9}} ydy$$

дорівнює нулю, як інтеграл у симетричних межах від непарної функції, отже $\mu_{xy} = 0$, тобто величини X та Y некорельовані.

Таким чином, з корельованості двох випадкових величин випливає їх залежність, але з залежності не випливає корельованість. Навпаки, з незалежності двох випадкових величин випливає їх некорельованість, але з некорельованості не випливає незалежність.

2.38. Двовимірний закон нормального розподілу

Означення. *Двовимірним нормальним законом розподілу* називається закон розподілу двовимірної величини (X,Y), якщо щільність розподілу має вигляд:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_{xy}^2}}\times$$

$$\times\exp\left(-\frac{1}{2\left(1-r_{xy}^2\right)}\left(\frac{(x-a_1)^2}{\sigma_x^2}+\frac{(y-a_2)^2}{\sigma_y^2}-2r_{xy}\frac{x-a_1}{\sigma_x}\cdot\frac{y-a_2}{\sigma_y}\right)\right). \qquad (*)$$

Поверхню в тривимірному просторі, що є «графіком» функції $f(x,y)$, наведено на рис. 39.

З формули (*) видно, що двовимірний нормальний закон розподілу характеризується п'ятьма параметрами: $a_1, a_2, \sigma_x, \sigma_y, r_{xy}$. Можна довести, що ці параметри мають наступний ймовірнісний зміст: a_1, a_2 – математичні сподівання відповідно компонент X та Y; σ_x, σ_y – їх середньоквадратичні відхилення; r_{xy} – коефіцієнт кореляції.

Теорема. *Якщо компоненти X та Y двовимірної нормально розподіленої випадкової величини (X,Y) некорельовані, то X та Y є незалежними.*

Доведення. Оскільки X та Y некорельовані, то $r_{xy} = 0$. Тоді щільність розподілу набуває вигляду:

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{1}{2}\left(\frac{(x-a_1)^2}{\sigma_x^2}+\frac{(y-a_2)^2}{\sigma_y^2}\right)\right)=$$

$$=\frac{1}{\sigma_x\sqrt{2\pi}}\exp\left(-\frac{(x-a_1)^2}{2\sigma_x^2}\right)\cdot\frac{1}{\sigma_y\sqrt{2\pi}}\exp\left(-\frac{(y-a_2)^2}{2\sigma_y^2}\right)=f_1(x)f_2(y).$$

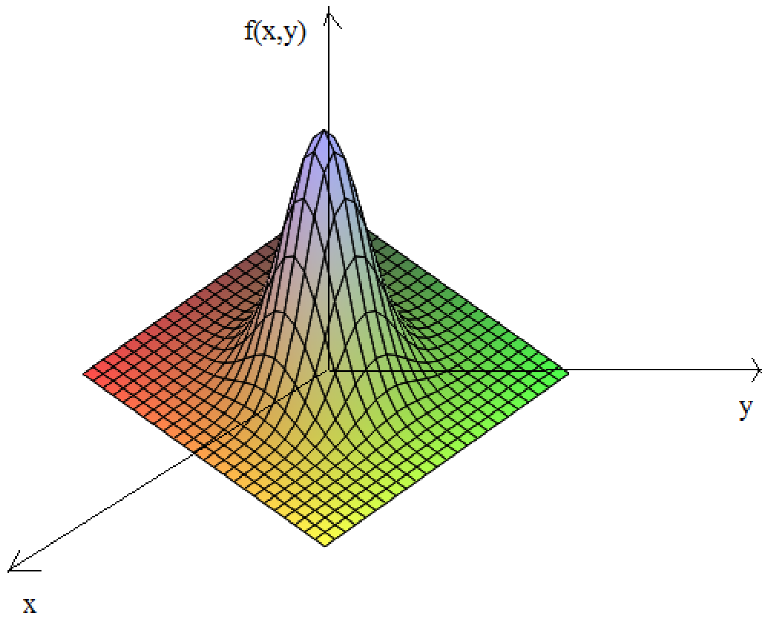


Рис. 39

Таким чином щільність сумісного розподілу системи (X,Y) дорівнює добутку щільностей розподілів величин X та Y, а це означає, що величини X та Y незалежні.

Наслідок. *Для того, щоб компоненти X та Y двовимірної нормально розподіленої випадкової величини були некорельовані, необхідно і достатньо, щоб величини X та Y були незалежні.*

2.39. Нерівність Чебишова

Ми знаємо, що завчасно передбачити, яке значення прийме випадкова величина в одному випробуванні, неможливо. Разом з цим, якщо ми маємо суму досить великого числа випадкових величин, то її поведінку при певних умовах можна передбачити. Більш того, в певному сенсі така сума втрачає випадковий характер і стає закономірною.

В теорії ймовірностей є низка теорем, де встановлюються умови, за яких сукупна дія дуже великого числа випадкових величин стає практично не випадковою, а закономірною. Ці теореми носять спільну назву *закону великих чисел*. До найбільш відомих з них відносяться теореми Чебишова та Бернуллі. Для доведення теореми Чебишова доведемо спочатку нерівність Чебишова[1].

Нерівність Чебишова. *Ймовірність того, що відхилення випадкової величини* X *від її математичного сподівання за абсолютною величиною буде меншим за додатне число* ε, *не менша, ніж* $1-D[X]/\varepsilon^2$:

$$P\left(\left|X-M[X]\right|<\varepsilon\right)\geq 1-\frac{D[X]}{\varepsilon^2} .$$

Доведення. Обмежимось випадком неперервних випадкових величин. Очевидно, що

$$P\left(\left|X-M[X]\right|<\varepsilon\right)=1-P\left(\left|X-M[X]\right|\geq\varepsilon\right), \qquad (*)$$

адже події $\left|X-M[X]\right|<\varepsilon$ та $\left|X-M[X]\right|\geq\varepsilon$ взаємно протилежні. Позначимо $m=M[X]$ і знайдемо:

$$D[X]=\int_{-\infty}^{+\infty}(x-m)^2 f(x)\,dx=\int_{-\infty}^{m-\varepsilon}(x-m)^2 f(x)\,dx+\int_{m-\varepsilon}^{m+\varepsilon}(x-m)^2 f(x)\,dx+$$

$$+\int_{m+\varepsilon}^{+\infty}(x-m)^2 f(x)\,dx\geq\int_{-\infty}^{m-\varepsilon}(x-m)^2 f(x)\,dx+\int_{m+\varepsilon}^{+\infty}(x-m)^2 f(x)\,dx .$$

Для значень x, що належать проміжку, за яким береться перший з останніх інтегралів, виконано $x\leq m-\varepsilon$, а для значень x, що належать проміжку, за яким береться другий з останніх інтегралів, виконано $x\geq m+\varepsilon$, тобто в обох випадках виконано $|x-m|\geq\varepsilon$, або $(x-m)^2\geq\varepsilon^2$. Отже:

$$\int_{-\infty}^{m-\varepsilon}(x-m)^2 f(x)\,dx\geq\varepsilon^2\int_{-\infty}^{m-\varepsilon}f(x)\,dx ,$$

$$\int_{m+\varepsilon}^{+\infty}(x-m)^2 f(x)\,dx\geq\varepsilon^2\int_{m+\varepsilon}^{+\infty}f(x)\,dx .$$

Звідси дістаємо:

$$D[X]\geq\varepsilon^2\left(\int_{-\infty}^{m-\varepsilon}f(x)\,dx+\int_{m+\varepsilon}^{+\infty}f(x)\,dx\right)=\varepsilon^2\left(P(X\leq m-\varepsilon)+P(X\geq m+\varepsilon)\right)=$$

$$=\varepsilon^2(P(X-m\leq-\varepsilon)+P(X-m\geq\varepsilon))=\varepsilon^2 P(|X-m|\geq\varepsilon),$$

а тоді:

$$P(|X-m|\geq\varepsilon)\leq\frac{D[X]}{\varepsilon^2} .$$

[1] Чебишов Пафнутій Львович (1821–1894) – видатний російський математик.

Враховуючи тепер (*), звідси отримуємо:

$$P\left(\left|X - M[X]\right| < \varepsilon\right) \geq 1 - \frac{D[X]}{\varepsilon^2} .$$

Зауваження. Нерівність Чебишова для практики має досить обмежене значення, оскільки дає грубу, а іноді навіть тривіальну оцінку. Наприклад, якщо $D[X] > \varepsilon^2$, то $D[X]/\varepsilon^2 > 1$, а отже $1 - D[X]/\varepsilon^2 < 0$, і тоді нерівність $P\left(\left|X - M[X]\right| < \varepsilon\right) \geq 1 - D[X]/\varepsilon^2$ вказує лише на те, що ймовірність відхилення невід'ємна, а це й так зрозуміло. Але теоретичне значення нерівності Чебишова досить велике, адже на його підставі можна довести інші, більш сильні твердження.

<u>Приклади</u>. 1. В освітлювальну мережу паралельно включено 20 ламп. Ймовірність того, що за проміжок часу T лампа буде включена, дорівнює 0,8. Користуючись нерівністю Чебишова, оцінити ймовірність того, що абсолютна величина різниці між числом включених ламп і середнім числом (математичним сподіванням) включених ламп за проміжок часу T менше, ніж 3.

Знайдемо математичне сподівання та дисперсію числа ламп, що включено, для чого скористаємось розподілом Бернуллі (див. п. 2.11):

$$M[X] = np = 20 \cdot 0,8 = 16, \quad D[X] = npq = 20 \cdot 0,8 \cdot 0,2 = 3,2 .$$

За нерівністю Чебишова (при $\varepsilon = 3$) маємо:

$$P(|X - M[X]| < 3) \geq 1 - \frac{D[X]}{3^2} = 1 - \frac{3,2}{9} \approx 0,64 .$$

2. Ймовірність появи події у кожному випробуванні дорівнює $1/4$. Користуючись нерівністю Чебишова, оцінити ймовірність того, що число X появ події міститься у межах від 150 до 250, якщо буде проведено 800 випробувань.

Знайдемо математичне сподівання та дисперсію величини X – числа появ події у 800 випробуваннях:

$$M[X] = np = 800 \cdot 0,25 = 200, \quad D[X] = npq = 800 \cdot 0,25 \cdot 0,75 = 150 .$$

Далі:

$$P(150 < X < 250) = P(-50 < X - 200 < 50) = P\left(|X - 200| < 50\right) \geq 1 - \frac{D[X]}{50^2} =$$

$$= 1 - \frac{150}{2500} = 0,94 .$$

3. Дискретну випадкову величину X задано законом розподілу:

X	0,1	0,4	0,6
P	0,2	0,3	0,5

Користуючись нерівністю Чебишова, оцінити $P\left(|X-M[X]|<\sqrt{0,4}\right)$.

Знайдемо:

$M[X]=0,1\cdot 0,2+0,4\cdot 0,3+0,6\cdot 0,5=0,44$,

$M[X^2]=0,1^2\cdot 0,2+0,4^2\cdot 0,3+0,6^2\cdot 0,5=0,23$,

$D[X]=M[X^2]-M^2[X]=0,23-0,44^2=0,0364$,

$$P\left(|X-M[X]|<\sqrt{0,4}\right)\geq 1-\frac{0,0364}{0,4}=0,909.$$

2.40. Теорема Чебишова (закон великих чисел)

Теорема Чебишова. *Якщо* $X_1, X_2, ..., X_n, ...$ *– попарно незалежні випадкові величини, які мають скінченні математичні сподівання, причому дисперсії цих величин рівномірно обмежені, тобто* $\forall n:\ D[X_n]\leq C$*, то для будь-якого числа* $\varepsilon>0$ *виконано граничну рівність:*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{M[X_1]+M[X_2]+\ldots+M[X_n]}{n}\right|<\varepsilon\right)=1.$$

Тобто якщо розглядається достатньо велика кількість незалежних випадкових величин, що мають обмежені дисперсії, то практично достовірною можна вважати подію, яка полягає в тому, що відхилення середнього арифметичного випадкових величин від середнього арифметичного їх математичних сподівань буде за абсолютною величиною як завгодно малим.

Доведення. Розглянемо наступну випадкову величину:

$$\overline{X}_n=\frac{X_1+X_2+\ldots+X_n}{n},$$

тобто середнє арифметичне величин $X_1, X_2, ..., X_n$. Знайдемо:

$$M[\overline{X}_n]=M\left[\frac{1}{n}(X_1+X_2+\ldots+X_n)\right]=\frac{1}{n}M[X_1+X_2+\ldots+X_n]=$$

$$=\frac{1}{n}\left(M[X_1]+M[X_2]+\ldots+M[X_n]\right).$$

Використовуючи нерівність Чебишова, дістанемо:

$$P\left(\left|\bar{X}_n - M[\bar{X}_n]\right| < \varepsilon\right) \geq 1 - \frac{D[\bar{X}_n]}{\varepsilon^2}.$$

Або:

$$P\left(\left|\bar{X}_n - \frac{1}{n}\left(M[X_1] + M[X_2] + \ldots + M[X_n]\right)\right| < \varepsilon\right) \geq 1 - \frac{D[\bar{X}_n]}{\varepsilon^2}.$$

За властивостями дисперсії можемо записати:

$$D[\bar{X}_n] = D\left[\frac{1}{n}(X_1 + X_2 + \ldots + X_n)\right] = \frac{1}{n^2}\left(D[X_1] + D[X_2] + \ldots + D[X_n]\right).$$

Оскільки $D[X_k] \leq C$, то:

$$D[\bar{X}_n] \leq \frac{1}{n^2}(C + C + \ldots + C) = \frac{nC}{n^2} = \frac{C}{n}.$$

Отже:

$$P\left(\left|\bar{X}_n - \frac{1}{n}\left(M[X_1] + M[X_2] + \ldots + M[X_n]\right)\right| < \varepsilon\right) \geq 1 - \frac{C}{n\varepsilon^2}.$$

Переходячи в цій рівності до границі при $n \to \infty$, отримаємо:

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \frac{1}{n}\left(M[X_1] + M[X_2] + \ldots + M[X_n]\right)\right| < \varepsilon\right) \geq 1.$$

Але, враховуючи те, що ймовірність події не може перевищувати одиниці, остаточно маємо:

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{M[X_1] + M[X_2] + \ldots + M[X_n]}{n}\right| < \varepsilon\right) = 1.$$

Теорему доведено.

Наслідок. *Якщо* $X_1, X_2, \ldots, X_n, \ldots$ *– попарно незалежні випадкові величини, що мають одне й те ж математичне сподівання* a*, та якщо дисперсії цих величин рівномірно обмежені, то для будь-якого числа* $\varepsilon > 0$ *справджується:*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - a\right| < \varepsilon\right) = 1.$$

Доведення. Дійсно, якщо $M[X_k] = a \;\; \forall\; k = 1, 2, \ldots$, то

$$\frac{M[X_1] + M[X_2] + \ldots + M[X_n]}{n} = \frac{a + a + \ldots + a}{n} = \frac{na}{n} = a,$$

і, користуючись теоремою Чебишова, отримуємо потрібне.

Теорема Чебишова стверджує, що, хоча незалежні випадкові величини можуть приймати значення, які значно відрізняються від своїх математичних сподівань, середнє арифметичне досить великого числа випадкових величин з великою ймовірністю приймає значення, близькі до певного сталого числа. Таким чином, це середнє з випадкового перетворюється фактично на закономірне.

Теорема Чебишова має велике значення для практики. Розглянемо такий приклад. Нехай ми проводимо вимірювання деякої фізичної величини. Як зазвичай, одним вимірюванням не обмежуються, а проводять декілька вимірювань і в якості шуканого значення приймають середнє арифметичне результатів вимірювань. За яких умов цей спосіб є виправданим? Відповідь на це питання дає теорема Чебишова (точніше, наслідок з неї). Дійсно, розглянемо результати всіх вимірювань як випадкові величини $X_1, X_2, ..., X_n$. До цих величин можна застосувати теорему Чебишова, якщо: 1) вони попарно незалежні; 2) мають одне й те ж математичне сподівання; 3) їх дисперсії рівномірно обмежені. Перша з цих умов виконуються, якщо результат кожного вимірювання не залежить від результатів інших вимірювань. Друга умова виконується, якщо вимірювання проведено без систематичних (одного знаку) похибок. У цьому випадку математичні сподівання всіх випадкових величин однакові і дорівнюють шуканому значенню a. Третя умова виконується, якщо прилад, яким проводяться вимірювання, забезпечує певну їх точність.

Якщо всі ці умови виконано, можемо застосувати теорему Чебишова. Відповідно неї ймовірність нерівності

$$\left|\frac{X_1 + X_2 + ... + X_n}{n} - a\right| < \varepsilon$$

при достатньо великих n стає як завгодно близькою до одиниці. Іншими словами, при достатньо великій кількості вимірювань майже достовірно, що середнє арифметичне їх результатів як завгодно мало відрізняється від шуканого значення величини, що вимірюється.

На теоремі Чебишова ґрунтується вибірковий метод у статистиці. Це питання розглядається у Розділі III.

Приклад. Послідовність незалежних випадкових величин $X_1, X_2, ..., X_n, ...$ задано законами їх розподілів:

Чи можна застосувати до цієї послідовності теорему Чебишова?

X_n	$-n\alpha$	0	$n\alpha$
P	$\frac{1}{2n^2}$	$1-\frac{1}{n^2}$	$\frac{1}{2n^2}$

Для того, щоб до послідовності величин можна було застосувати теорему Чебишова, достатньо, щоб ці величини були попарно незалежні, мали скінченні математичні сподівання та рівномірно обмежені дисперсії.

Оскільки дані величини незалежні, то вони й попарно незалежні, тобто першу умову теореми Чебишова виконано. Розглянемо математичні сподівання:

$$M[X_n] = -n\alpha \cdot \frac{1}{2n^2} + 0 \cdot \left(1 - \frac{1}{n^2}\right) + n\alpha \cdot \frac{1}{2n^2} = 0 .$$

Таким чином, другу умову також виконано. Знайдемо дисперсії:

$$D[X_n] = M[X_n^2] - M^2[X_n] = M[X_n^2] =$$

$$= n^2\alpha^2 \cdot \frac{1}{2n^2} + 0 \cdot \left(1 - \frac{1}{n^2}\right) + n^2\alpha^2 \cdot \frac{1}{2n^2} = \alpha^2 .$$

Таким чином, всі дисперсії рівномірно обмежені числом α^2, і третю умову також виконано.

Отже, до даної послідовності можна застосувати теорему Чебишова.

2.41. Теорема Бернуллі

Важливим наслідком теореми Чебишова є теорема Бернуллі, яка встановлює зв'язок між відносною частотою та ймовірністю події (див. п. 1.6). Цю теорему було доведено у 1713 році Я. Бернуллі (вона вперше й отримала назву закону великих чисел), а у 1846 році П. Л. Чебишов довів свою теорему, з якої теорема Бернуллі випливає як наслідок.

Теорема Бернуллі. *Якщо у кожному з n незалежних випробувань ймовірність p появи події A є сталою, то справджується гранична рівність:*

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1,$$

де m – число появ події A в n випробуваннях.

Інакше кажучи, ймовірність того, що відхилення відносної частоти m/n від ймовірності p події за абсолютною величиною буде скільки завгодно малим, наближається до одиниці зі зростанням числа n випробувань.

Доведення. Позначимо як X_1 – ДВВ, яка дорівнює числу появ події A у 1-му випробуванні; як X_2 – ДВВ, яка дорівнює числу появ події A у 2-му випробуванні; … ; як X_n – числу появ події A у n -му випробуванні.

Закон розподілу кожної з величин X_k $(k = 1, 2, ..., n)$ має вигляд (див. п. 2.11):

X_k	0	1
P	q	p

Тут $q = 1 - p$. Перевіримо можливість застосування до величин $X_1, X_2, ..., X_n$ теореми Чебишова. По-перше, треба, щоб ці величини були попарно незалежні. Це випливає з того, що незалежними є випробування. По-друге, треба, щоб вони мали скінченні математичні сподівання. Оскільки $M[X_k] = p$ $(k = 1, 2, ..., n)$, то цю умову виконано. По-третє, треба, щоб вони мали рівномірно обмежені дисперсії. Розглянемо (див. п. 2.11):

$$D[X_k] = pq = p(1-p).$$

Знайдемо найбільше значення функції $\varphi(p) = p(1-p)$ на відрізку $[0,1]$. Маємо: $\varphi(0) = \varphi(1) = 0$; похідна $\varphi'(p) = 1 - 2p$ перетворюється на нуль в точці $p = 1/2$; $\varphi(1/2) = 1/4$. Отже, $\max\limits_{[0,1]} \varphi(p) = 1/4$, і таким чином дисперсії всіх величин $X_1, X_2, ..., X_n$ не перевищують числа $1/4$. Таким чином, третю умову також виконано, і ми можемо до величин $X_1, X_2, ..., X_n$ застосувати теорему Чебишова.

Оскільки всі величини $X_1, X_2, ..., X_n$ мають одне й те ж математичне сподівання p, то наслідок з теореми Чебишова (див. п. 2.40) дає:

$$\lim_{n \to \infty} P\left(\left| \frac{X_1 + X_2 + ... + X_n}{n} - p \right| < \varepsilon \right) = 1.$$

Покажемо, що $X_1 + X_2 + ... + X_n = m$. Дійсно, величина X_k $(k = 1, 2, ..., n)$ дорівнює числу появ події A у k-му випробуванні, отже, сума $X_1 + X_2 + ... + X_n$ дорівнює числу появ події A у всіх n випробуваннях, тобто саме m. Отже:

$$\frac{X_1 + X_2 + ... + X_n}{n} = \frac{m}{n}.$$

Враховуючи цю рівність, остаточно дістанемо:

$$\lim_{n \to \infty} P\left(\left| \frac{m}{n} - p \right| < \varepsilon \right) = 1,$$

що й треба було довести.

Контрольні питання до розділу II

1. Що таке випадкова величина? Які є види випадкових величин?
2. Що таке дискретна випадкова величина? Як вона задається?
3. Що таке математичне сподівання дискретної випадкової величини? Які його основні властивості та ймовірнісний зміст?
4. Що таке дисперсія дискретної випадкової величини? Які її основні властивості та ймовірнісний зміст?
5. Які особливості числових характеристик однаково розподілених взаємно незалежних випадкових величин?
6. Що таке початкові та центральні теоретичні моменти? Для чого вони використовуються?
7. Яка величина називається розподіленою за біноміальним розподілом? Чому дорівнюють математичне сподівання та дисперсія такої величини?
8. Яка величина називається розподіленою за законом Пуассона? Чому дорівнюють математичне сподівання та дисперсія такої величини?
9. Що таке неперервна випадкова величина? Що таке функція розподілу неперервної випадкової величини? Чи мають дискретні випадкові величини функцію розподілу?
10. Що таке щільність розподілу неперервної випадкової величини? У чому полягає її ймовірнісний зміст? Чи мають дискретні випадкові величини щільність розподілу?
11. Як визначаються математичне сподівання та дисперсія неперервної випадкової величини?
12. Яка величина називається розподіленою рівномірно на заданому відрізку? Чому дорівнюють математичне сподівання та дисперсія такої величини?
13. Яка величина називається розподіленою за показниковим законом? Чому дорівнюють математичне сподівання та дисперсія такої величини?
14. Яка величина називається розподіленою за нормальним законом? Чому дорівнюють математичне сподівання та дисперсія такої величини?
15. Як впливають параметри нормального розподілу на форму кривої Гауса?
16. У чому полягає правило «трьох сигм»?
17. Як оцінюються відхилення заданого розподілу від нормального?
18. У чому суть центральної граничної теореми Ляпунова?
19. Що таке двовимірні випадкові величини? Як задаються двовимірні дискретні випадкові величини?
20. Що таке функція розподілу двовимірної випадкової величини?
21. Що таке щільність розподілу двовимірної випадкової величини?
22. Які випадкові величини називаються незалежними?

23. Які характеристики системи двох випадкових величин характеризують тісноту зв'язку між ними?
24. Який зв'язок між корельованістю та залежністю випадкових величин?
25. У чому полягає суть закону великих чисел?

Вправи для самостійного розв'язування

1. Серед 10 деталей 4 нестандартних. Навмання відбираються 3 деталі. Написати ряд розподілу випадкової величини – кількості нестандартних деталей серед 4-х відібраних.

2. На шляху руху автомобіля 6 світлофорів, кожен з яких дозволяє або забороняє рух автомобіля з ймовірністю 0,5. Скласти ряд розподілу і побудувати функцію розподілу кількості світлофорів, які автомобіль минув без зупинки. Чому дорівнює математичне сподівання та дисперсія цієї випадкової величини?

3. Знайти математичне сподівання дискретної випадкової величини X – числа таких кидань п'яти гральних кісток, у кожному з яких на двох кістках з'явиться по одному очку, якщо загальне число кидань дорівнює 20.

4. Знайти закон розподілу дискретної випадкової величини X, яка може приймати лише два значення: x_1 з ймовірністю 0,6 та x_2 $(x_1 < x_2)$, якщо $M[X]=1{,}4$; $D[X]=0{,}24$.

5. Знайти дисперсію дискретної випадкової величини X – числа появ події A в двох незалежних випробуваннях, якщо ймовірність появи події в кожному випробуванні однакова, і відомо, що $M[X]=1{,}2$.

6. Проводяться незалежні випробування з однаковою ймовірністю появи події A в кожному випробуванні. Знайти ймовірність появи події A, якщо дисперсія числа появ події A в трьох незалежних випробуваннях дорівнює 0,63.

7. Задано функцію розподілу неперервної випадкової величини X:

$$F(x)=\begin{cases} 0, \; x \le -1, \\ ax+b, \; -1 < x \le \dfrac{1}{3}, \\ 1, \; x > \dfrac{1}{3}. \end{cases}$$

Знайти параметри a і b, а також ймовірність того, що внаслідок випробування величина X прийме значення з проміжку $(0,1/3)$.

8. Для величини X з попередньої задачі знайти щільність розподілу. Побудувати графіки функції та щільності розподілу.

9. Задано щільність розподілу неперервної випадкової величини X:

$$f(x)=\begin{cases} 0, \; x \le 0, \\ A\cos x, \; 0 < x \le \pi/2, \\ 0, \; x > \pi/2. \end{cases}$$

Знайти параметр A і функцію розподілу величини X. Побудувати графіки функції та щільності розподілу.

10. Задано щільність розподілу випадкової величини X:

$$f(x)=\begin{cases}\dfrac{A}{\sqrt{4-x^2}}, & x\in(-2,2),\\ 0, & x\notin(-2,2).\end{cases}$$

Знайти

1) параметр A;

2) функцію розподілу і побудувати графіки функції та щільності;

3) ймовірність того, що внаслідок випробування величина X прийме значення з інтервалу $(1,3)$;

4) математичне сподівання, дисперсію та середньоквадратичне відхилення величини X.

11. Задано щільність розподілу випадкової величини X:

$$f(x)=Ae^{-|x|},\ -\infty<x<+\infty.$$

Знайти коефіцієнт A, математичне сподівання та дисперсію величини X.

12. Задано щільність розподілу випадкової величини X:

$$f(x)=\frac{A}{e^x+e^{-x}},\ -\infty<x<+\infty.$$

Знайти

1) коефіцієнт A;

2) функцію розподілу величини X;

3) ймовірність того, що в двох незалежних випробуваннях величина X прийме значення, менші за 1.

13. Вимірювальний прилад має систематичну похибку 5 м та середньоквадратичну похибку 75 м. Знайти ймовірність того, що похибка вимірювання не перевищить за абсолютною величиною 5 м, якщо вважати, що вона розподілена за нормальним законом.

14. Випадкова величина X має нормальний розподіл, причому $a=0$, $\sigma=1$. Що більше: $P(-0{,}5\le X\le -0{,}1)$ або $P(1\le X\le 2)$?

15. Зважують деяку речовину без систематичних (тобто одного знаку) похибок. Випадкові похибки підпорядковані нормальному закону з середньоквадратичним відхиленням $\sigma=20$ г. Знайти ймовірність того, що зважування буде виконано з похибкою, яка за абсолютною величиною на перевищує 10 г.

16. Випадкова величина X розподілена нормально з параметром $a=10$. Ймовірність потрапляння величини X в інтервал (10, 20) дорівнює 0,3. Чому дорівнює ймовірність потрапляння величини X в інтервал (0, 10)?

17. Задано закон розподілу двовимірної випадкової величини:

Y	X		
4	3	10	12
5	0,17 0,10	0,13 0,30	0,25 0,05

Знайти закони розподілу величин X та Y.

18. Задано функцію розподілу двовимірної випадкової величини:

$$F(x,y)=\begin{cases}\sin x\cdot\sin y,\ 0\le x\le \pi/2,\ 0\le y\le\pi/2,\\ \quad 0,\ \textit{якщо}\ x<0\ \textit{або}\ \ y<0.\end{cases}$$

Знайти ймовірність потрапляння випадкової точки (X,Y) в прямокутник, обмежений прямими $x=0$, $x=\pi/4$, $y=\pi/6$, $y=\pi/3$.

19. Щільність $f(x,y)$ розподілу двовимірної випадкової величини дорівнює $A\left(R-\sqrt{x^2+y^2}\right)$ в крузі $x^2+y^2\le R^2$. Зовні цього круга $f(x,y)=0$. Знайти:

1) параметр A;

2) ймовірність потрапляння випадкової точки (X,Y) в круг радіусу $r<R$ з центром у початку координат.

20. Двовимірну випадкову величину (X,Y) задано щільністю розподілу:

$$f(x,y)=\begin{cases}A,\ \textit{якщо}\,|x|\le 1,\ |y|\le 2,\\ 0,\ \textit{якщо}\,|x|>1\ \ \textit{або}\ \ |y|>2.\end{cases}$$

Знайти:

1) параметр A;

2) щільності розподілів складових X та Y.

21. Користуючись нерівністю Чебишова, оцінити ймовірність того, що будь-яка випадкова величина X відхилиться від свого математичного сподівання менш, ніж на три середньоквадратичних відхилення цієї величини.

22. Дискретну випадкову величину X задано законом розподілу:

X	0,3	0,6
P	0,2	0,8

Користуючись нерівністю Чебишова, оцінити ймовірність того, що $|X-M[X]|<0{,}2$.

23. Прилад складається з 10 незалежно працюючих елементів. Для кожного елемента ймовірність його відмови за час T дорівнює 0,05. Користуючись нерівністю Чебишова, оцінити ймовірність того, що абсолютна величина різни-

ці між числом елементів, що відмовили, і середнім числом (математичним сподіванням) відмов за час T буде менша ніж 2.

24. Дано послідовність незалежних випадкових величин $X_1, X_2, ..., X_n, ...$. Випадкова величина X_n має закон розподілу:

X_n	$-\sqrt{n}$	0	$\sqrt{n}$
P	$1/n$	$1-2/n$	$1/n$

Чи можна застосувати до цієї послідовності теорему Чебишова?

25. Дано послідовність незалежних випадкових величин $X_1, X_2, ..., X_n, ...$. Випадкова величина X_n має закон розподілу:

X_n	$-a$	a
P	$n/(2n+1)$	$(n+1)/(2n+1)$

Чи можна застосувати до цієї послідовності теорему Чебишова?

Розділ III. Елементи математичної статистики

3.1. Предмет і задачі математичної статистики
Короткий історичний нарис

Математична статистика (МС) – це наука, яка вивчає способи відбору та обробки статистичних даних для одержання наукових та практичних висновків. Ці дані носять не абсолютний, а ймовірнісний характер, тому їх аналіз вимагає широкого використання методів теорії ймовірностей. Таким чином, МС може розглядатися як застосування методів теорії ймовірностей до задач практики.

Як і теорія ймовірностей, МС це наука суттєво математична, і використовує різноманітний математичний апарат.

До основних задач МС відносяться наступні.

1. Вказання способів відбору та групування статистичних відомостей.

2. Розробка методів аналізу статистичних даних в залежності від мети дослідження. Сюди відносяться:

а) оцінка невідомої ймовірності події; оцінка параметрів розподілу, вигляд якого відомо; оцінка залежності випадкової величини від однієї або декількох інших випадкових величин;

б) перевірка статистичних гіпотез про вигляд невідомого розподілу.

Сучасна МС розробляє способи визначення числа необхідних випробувань до початку дослідження (планування експерименту), в процесі дослідження (послідовний аналіз), та вирішує багато інших задач. Тому сучасну МС визначають як науку про прийняття рішень в умовах невизначеності.

МС виникла у XVII ст. і розвивалася паралельно з теорією ймовірностей. Подальшим розвитком МС зобов'язана таким вченим як П. Л. Чебишов, А. А. Марков, О. М. Ляпунов, А. Кельтон, К. Пірсон, Р. Стьюдент, Р. Фішер та іншим.

Методи МС широко використовуються і в задачах геології. Вперше це відбулося в працях Ч. Лайєля (1833 р.), в яких були використані найпростіші способи статистичної обробки палеонтологічних даних. У подальшому методи МС використовуються також в літології, розвідці корисних копалин. Систематичне використання в геології математичних методів, зокрема, методів МС, починається у 40-і роки XX ст. з робіт А. Вістеліуса. Особливо інтенсивний процес проникнення математичних методів в геологію розпочався з появою та розвитком електронно-обчислювальної техніки, комп'ютерів.

В сучасних геологічних дослідженнях використовуються всі напрями математики – математичний аналіз, диференціальні рівняння, лінійна алгебра, теорія множин, теорія груп, геометрія, теорія ймовірностей і математична статистика та багато інших. Зокрема, апарат теорії ймовірностей та МС використовується при вивченні розподілів геологічних ознак та їх застосуванні при розв’язанні геологічних задач, перевірці статистичних гіпотез про однорідність сукупностей, застосуванні методів кореляції та регресії та ін.

3.2. Вибірковий метод. Генеральна та вибіркова сукупності

Основним методом МС є так званий *вибірковий метод*. Полягає він в наступному. Нехай потрібно визначити сукупність об’єктів відносно деякої якісної або кількісної ознаки, яка характеризує ці об’єкти. Кожен об’єкт, що спостерігається, має декілька ознак. Розглядаючи лише одну ознаку кожного об’єкта, ми припускаємо, що інші ознаки рівноправні, або що сукупність об’єктів *однорідна*. Наприклад, якщо досліджують партію деталей, то якісною ознакою може бути стандартність або нестандартність кожної деталі, а кількісною ознакою – розмір деталі. Кількісні ознаки можуть бути неперервними та дискретними.

Іноді проводять повне обстеження всієї сукупності, тобто перевіряється кожен з об’єктів сукупності відносно ознаки, яка цікавить. Але практично таке обстеження недоцільне або, навіть, практично неможливе. Наприклад, коли кількість об’єктів сукупності занадто велика, або перевірка всіх об’єктів пов’язана з великими матеріальними витратами. Тому частіше обстежують тільки певну частину об’єктів, яку випадково відбирають зі всієї сукупності.

Генеральною сукупністю називають всю сукупність об’єктів, яка підлягає обстеженню.

Вибірковою сукупністю (*вибіркою*) називають сукупність об’єктів, які випадковим чином відбираються для обстеження з генеральної сукупності.

Об’ємом сукупності (генеральної або випадкової) називається число об’єктів цієї сукупності.

Наприклад, якщо з 5000 виробів для обстеження взято 50, то об’єм генеральної сукупності $N = 5000$, а об’єм вибірки $n = 50$.

Складати вибірку можна двома способами: після того, як об’єкт відібрано з сукупності та обстежено, його може бути повернено до генеральної сукупності, а може бути не повернено. Відповідно цьому вибірки розділяються на повторні та безповторні.

Повторною називають вибірку, при якій відібраний об’єкт повертається до генеральної сукупності перед відбором наступного об’єкта.

Безповторною називають вибірку, при якій відібраний об’єкт до генеральної сукупності не повертається.

Найчастіше використовують саме безповторні вибірки.

Вибірку можна ефективно використовувати для вивчення відповідної ознаки генеральної сукупності лише тоді, коли дані вибірки вірно відображають цю ознаку. Саме тоді вибірковий метод є виправданим. Коротко ця умова формулюється так: вибірка має бути *репрезентативною* (або *представницькою*).

Згідно з законом великих чисел (див. п. 2.40) можна стверджувати, що вибірка буде репрезентативною, якщо виконуються наступні умови:

1) вибірка здійснюється випадково: кожен об'єкт вибірки відібрано випадково з генеральної сукупності, та всі об'єкти генеральної сукупності мають однакову ймовірність потрапляння до вибірки;
2) об'єм вібірки, хоча значно менший, ніж об'єм генеральної сукупності, все ж таки достатньо великий для забезпечення дії закону великих чисел.

3.3. Способи відбору

У практичній діяльності використовують різноманітні способи відбору об'єктів з генеральної сукупності. Усі способи можна поділити на два види.

1. Відбір, який не вимагає розділення генеральної сукупності на частини. До цього виду відбору відносять:

а) простий випадковий безповторний відбір;

б) простий випадковий повторний відбір.

2. Відбір, при якому генеральна сукупність розділяється на частини (розшарований випадковий відбір). До цього виду відбору відносять:

а) типовий відбір;

б) механічний відбір;

в) серійний відбір.

Простим випадковим називають відбір, при якому об'єкти відбираються по одному зі всієї генеральної сукупності. Якщо об'єм генеральної сукупності великий, такий відбір дуже складний.

Типовим називають відбір, при якому об'єкти відбирають не з усієї генеральної сукупності, а лише з її типових частин. Наприклад, якщо вироби виготовлені на різних станках, то відбір проводять лише з виробів кожного станка окремо.

Типовий відбір доцільно використовувати тоді, коли ознака, що обстежується, помітно розрізняється в різних типових частинах генеральної сукупності.

Механічним називають відбір, при якому генеральна сукупність механічно поділяється на стільки частин, скільки має бути об'єктів у вибірці. А потім з кожної частини випадковим чином відбирають один об'єкт.

*Серійни*м називають відбір, при якому об'єкти з генеральної сукупності відбирають не по одному, а серіями, які обстежуються. Серійний відбір використовують тоді, коли ознака, яку досліджують, мало змінюється в різних серіях.

На практиці також використовують *комбінований* відбір. Наприклад, спочатку поділяють генеральну сукупність на серії однакового об'єму, потім випадковим чином відбирають декілька серій і, нарешті, з кожної серії випадковим чином беруть окремі об'єкти.

3.4. Статистичний розподіл вибірки

Нехай з генеральної сукупності дістали вибірку, причому значення x_1 ознаки спостерігалося n_1 разів, значення x_2 – n_2 разів, . . . , значення x_k – n_k разів. $n_1 + n_2 + \ldots + n_k = n$ – об'єм вибірки. Значення $x_1, x_2, \ldots, x_k$, що спостерігаються, називаються *варіантами*. Послідовність варіант, які записано у порядку зростання, називають *варіаційним рядом*. Числа $n_1, n_2, ..., n_k$ називаються *частотами*, а числа $w_1 = n_1/n, w_2 = n_2/n, ..., w_k = n_k/n$ – *відносними частотами*. Сума відносних частот: $w_1 + w_2 + ... + w_k = 1$.

Статистичним розподілом вибірки називається перелік варіант з відповідними їм частотами, або відносними частотами. Тобто таблиця наступного вигляду:

x_i	x_1	x_2	$\cdots$	x_k
n_i	n_1	n_2	$\cdots$	n_k

Або

x_i	x_1	x_2	$\cdots$	n_k
w_i	w_1	w_2	$\cdots$	w_k

Можна сказати, що статистичний розподіл відносних частот вибірки є статистичним аналогом закону розподілу випадкової величини.

Приклад. Нехай у вибірці значення $x_1 = 4$ зустрілося 5 разів, значення $x_2 = 7$ – 2 рази, значення $x_3 = 8$ – 3 рази, значення $x_4 = 12$ – 10 разів. Побудувати статистичний розподіл частот, а також розподіл відносних частот.

Статистичний розподіл частот має вигляд:

x_i	4	7	8	12
n_i	5	2	3	10

Об'єм вибірки: 5+2+3+10=20.

Відносні частоти: $w_1 = 5/20 = 0,25$, $w_2 = 2/20 = 0,10$, $w_3 = 3/20 = 0,15$, $w_4 = 10/20 = 0,50$.

Статистичний ряд відносних частот такий:

x_i	4	7	8	12
w_i	0,25	0,10	0,15	0,50

Контроль правильності: 0,25+0,10+0,15+0,50=1.

3.5. Емпірічна функція розподілу

Нехай є статистичний розподіл частот деякої ознаки X. Позначимо як n загальну кількість спостережень. Нехай x – довільне дійсне число (точка числової прямої). Позначимо як n_x – кількість спостережень, при яких спостерігалися ознаки X, що менше x.

Тоді відносна частота події $X < x$ дорівнює n_x/n. Якщо x змінюється, то, взагалі кажучи, змінюється й n_x, а отже й n_x/n, тобто n_x/n є функцією від x. Ця функція знаходиться емпіричним (дослідним) шляхом, тому її називають емпіричною.

Означення. *Емпірічною функцією розподілу* називається функція $F_n(x)$, яка визначає для кожного значення x відносну частоту події $X < x$.

Тобто за означенням:

$$F_n(x) = \frac{n_x}{n}.$$

Означення. Інтегральну функцію розподілу $F(x)$ генеральної сукупності називають *теоретичною функцію розподілу*.

Функція $F(x)$ відрізняється від функції $F_n(x)$ тим, що вона визначає ймовірність події $X < x$, а функція $F_n(x)$ – відносну частоту цієї події.

Якщо до функцій $F(x)$ та $F_n(x)$ застосувати теорему Бернуллі (див. п. 2.41), то отримаємо:

$$\lim_{n\to\infty} P(|F_n(x) - F(x)| < \varepsilon) = 1$$

для будь-якого $\varepsilon > 0$. Тобто при збільшенні числа випробувань n подія, яка полягає в тому, що емпірична функція розподілу буде мало відрізнятися від теоретичної функції розподілу, стає практично достовірною. А звідси випливає доцільність використання функції $F_n(x)$ для наближеного зображення теоретичної функції розподілу $F(x)$ генеральної сукупності. Можна сказати, що функція $F_n(x)$ є статистичним аналогом функції $F(x)$.

Такий висновок підтверджується також тим, що функція $F_n(x)$ має властивості, які аналогічні властивостям функції $F(x)$. А саме.

1. Всі значення функції $F_n(x)$ належать відрізку $[0,1]$:
$0 \le F_n(x) \le 1$.
2. Функція $F_n(x)$ є неспадною, тобто, якщо $x_1 < x_2$, то $F_n(x_1) \le F_n(x_2)$.
3. $\lim_{x\to-\infty} F_n(x) = 0,\ \lim_{x\to+\infty} F_n(x) = 1$.

Ця властивість фактично означає наступне: якщо x_1 – найменша варіанта, то $F_n(x) = 0$ при $x \le x_1$; якщо x_k – найбільша варіанта, то $F_n(x) = 1$ при $x > x_k$.

Приклад. Знайти емпіричну функцію розподілу за даним статистичним розподілом вибірки:

x_i	2	6	10
n_i	12	18	30

Побудувати графік функції $F_n(x)$.

Об'єм вибірки дорівнює $n = 12 + 18 + 30 = 60$. Найменша варіанта $x_1 = 2$, отже, $F_n(x) = 0$ при $x \le 2$. Найбільша варіанта $x_3 = 10$, отже, $F_n(x) = 1$ при $x > 10$. Знайдемо $F_n(x)$ при $2 < x \le 6$. Подія $X < x$ у цьому випадку еквівалентна події $X = 2$, а відносна частота цієї події $12/60 = 0{,}2$, тобто $F_n(x) = 0{,}2$. При $6 < x \le 10$ подія $X < x$ еквівалентна сумі подій $X = 2$, $X = 6$, і відносна частота цієї суми $(12 + 18)/60 = 30/60 = 0{,}5$, тобто $F_n(x) = 0{,}5$. Отже, дістаємо:

$$F_n(x) = \begin{cases} 0, & x \le 2, \\ 0{,}2, & 2 < x \le 6, \\ 0{,}5, & 6 < x \le 10, \\ 1, & x > 10. \end{cases}$$

Графік цієї функції має вигляд:

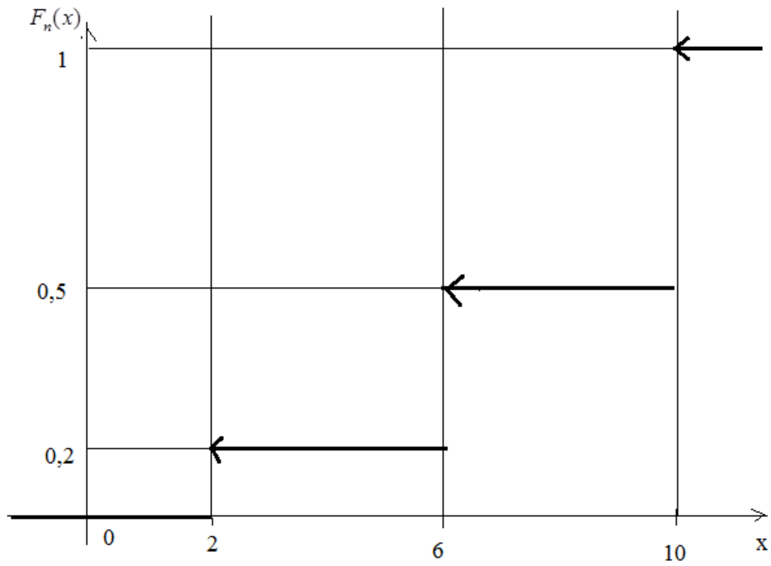


Рис. 40

Цей графік можна розглядати як наближений графік теоретичної функції розподілу. Він демонструє перелічені вище властивості функції $F_n(x)$, а також наступні властивості: ця функція кусково-стала, у точках, що відповідають значенням ознаки, вона має розриви I роду і неперервна зліва в цих точках. Легко помітити, що ці властивості аналогічні властивостям функції розподілу дискретної випадкової величини (див. п. 2.16).

3.6. Полігон і гістограма

З метою наглядного відображення даних вибірки використовуються графічні засоби. До них, зокрема, відносяться полігон і гістограма.

Нехай задано вибірку:

x_i	x_1	x_2	$\cdots$	x_k
n_i	n_1	n_2	$\cdots$	n_k

$n = n_1 + n_2 + \ldots + n_k$ – об’єм вибірки.

Означення. *Полігоном частот* вибірки називається ламана лінія на координатній площині, вершини якої знаходяться у точках (x_1, n_1), (x_2, n_2), … , (x_k, n_k).

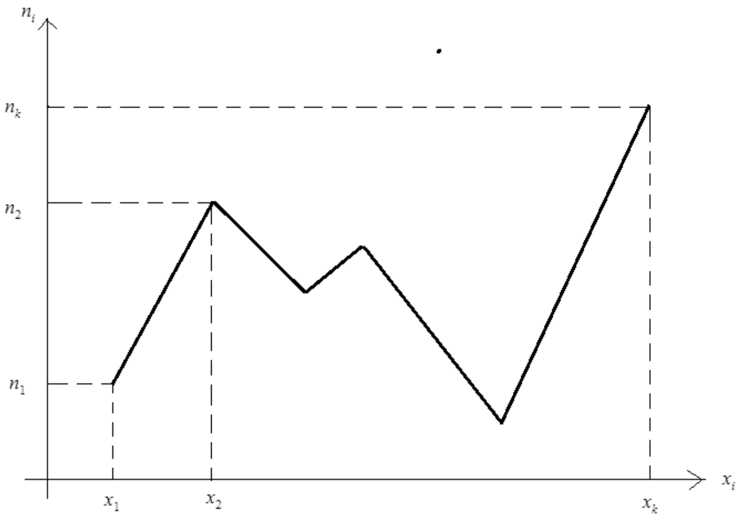


Рис. 41

Означення. *Полігоном відносних частот* вибірки називається ламана лінія на координатній площині, вершини якої знаходяться у точках (x_1, w_1), (x_2, w_2), … , (x_k, w_k), де $w_j = n_j / n \quad (j = 1, 2, ..., k)$ – відносні частоти вибірки.

Приклад. Нехай задано вибірку:

x_i	1	2	4	6	7
n_i	10	20	15	5	30

Побудувати полігони частот та відносних частот.

Полігон частот має вигляд:

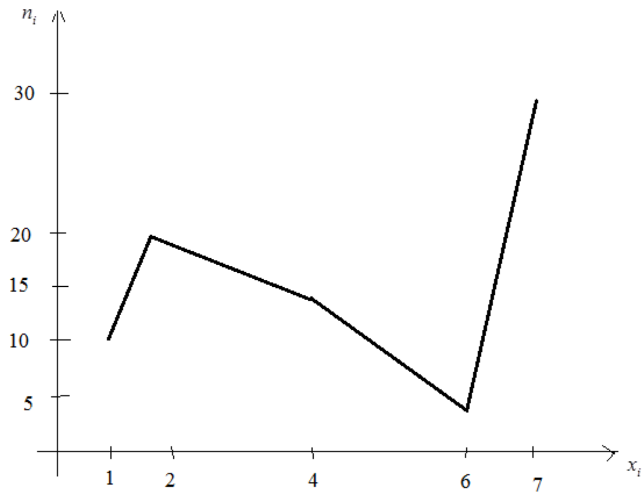


Рис. 42

Знайдемо відносні частоти:

$n = 10 + 20 + 15 + 5 + 30 = 80$,

$w_1 = \frac{10}{80} = 0{,}125$, $w_2 = \frac{20}{80}$, $w_3 = \frac{15}{80} = 0{,}1875$, $w_4 = \frac{5}{80} = 0{,}0625$,

$w_5 = \frac{30}{80} = 0{,}375$.

Статистичний розподіл відносних частот:

x_i	1	2	4	6	7
w_i	0,125	0,25	0,1875	0,0625	0,375

Полігон відносних частот має вигляд:

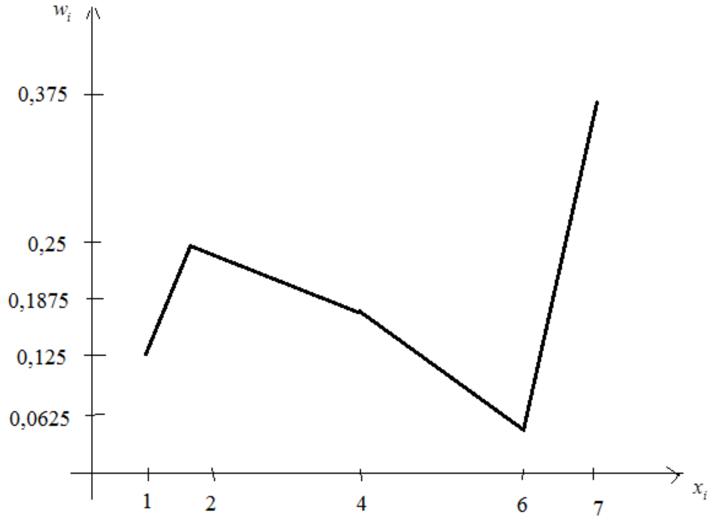


Рис. 43

Якщо ознака неперервна, або число різних елементів вибірки досить велике, то замість полігону використовують *гістограму*. Будується вона наступним чином. Розглянемо деякий інтервал (a,b), до якого належать всі елементи вибірки. Розіб'ємо цей інтервал на N частинних проміжків. Для спрощення будемо вважати, що всі ці проміжки мають однакову довжину $h = (b-a)/N$, хоча у загальному випадку довжини цих проміжків можуть бути різними. Для кожного частинного проміжку знайдемо суму частот тих елементів вибірки, які потрапили до цього проміжку. Позначимо як n_i^* – суму частот елементів вибірки, які потрапили до i-го проміжку $(i = 1, 2, \ldots, N)$. Тоді $n_1^* + n_2^* + \ldots + n_N^* = n$. На кожному частинному проміжку побудуємо прямокутника, основа якого співпадає з частинним проміжком, а висота дорівнює n_i^*/h, де i – номер проміжку. Висоти прямокутників, таким чином, пропорційні кількості елементів вибірки, які потрапили у відповідний проміжок – чим більше елементів потрапило у частинний проміжок, тим вище прямокутник, побудований на цьому проміжку. Тобто ці висоти відображають щільність вмісту елементів вибірки у відповідному проміжку.

Сукупність цих прямокутників є ступінчатою фігурою, яка й називається *гістограмою*.

Площа i-го прямокутника дорівнює $\left(n_i^*/h\right) \cdot h = n_i^*$, тобто кількості елементів, що потрапили до i-го проміжку. Таким чином, площа всієї гістограми дорівнює об'єму вибірки.

Іноді будують гістограму відносних частот. Вона відрізняється тим, що висота i-го прямокутника дорівнює $n_i^*/(hn)$. Таким чином, його площа n_i^*/n, а площа всієї гістограми:

$$\frac{n_1^*}{n} + \frac{n_2^*}{n} + \ldots + \frac{n_N^*}{n} = \frac{n_1^* + n_2^* + \ldots + n_N^*}{n} = \frac{n}{n} = 1.$$

Гістограма, таким чином, є статистичним аналогом щільності розподілу випадкової величини, її контур наближено відображає графік відповідної щільності розподілу. Наприклад, якщо висоти прямокутників приблизно однакові, то це свідчить про те, що розподіл відповідної випадкової величини близький до рівномірного (див. п. 2.21).

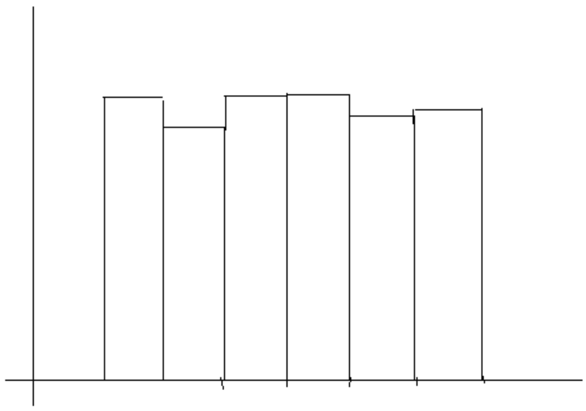


Рис. 44

Якщо висоти прямокутників помітно спадають з ростом номеру частинного проміжку, то це свідчить про те, що розподіл відповідної випадкової величини близький до показникового (див. п. 2.22).

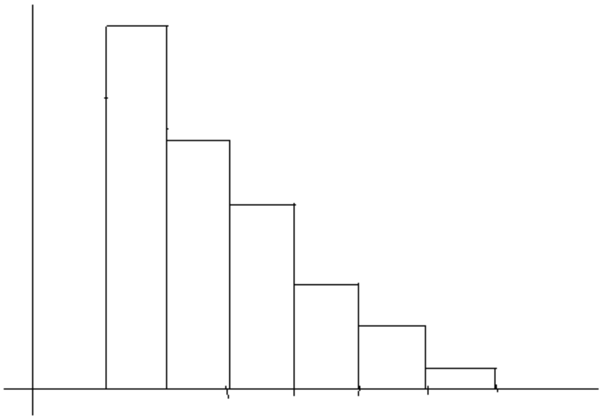


Рис. 45

Якщо висоти прямокутників, які знаходяться ближче до середини інтервалу (a,b), помітно більше, ніж висоти прямокутників, що знаходяться ближче до кінців інтервалу (a,b), то це свідчить про те, що розподіл близький до нормального (див. п. 2.25).

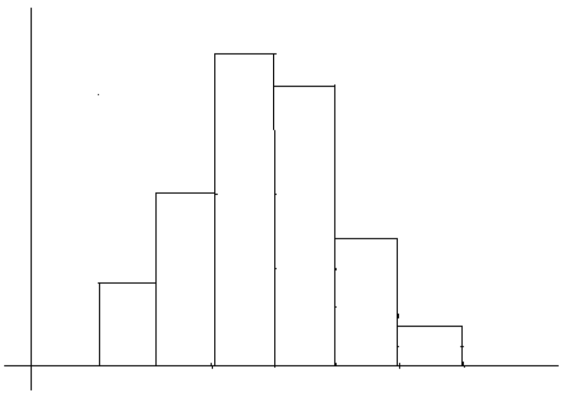


Рис. 46

<u>Приклад</u>. Нехай відомо наступний розподіл частот вибірки:

Частинні проміжки $(h=5)$	Сума частот вибірки n_i^*	Щільність частот n_i^*/h
5 – 10	4	0,8
10 – 15	6	1,2
15 – 20	16	3,2
20 – 25	36	7,2
25 –30	24	4,8
30 – 35	10	2,0
35 – 40	4	0,8

Гістограма має вигляд:

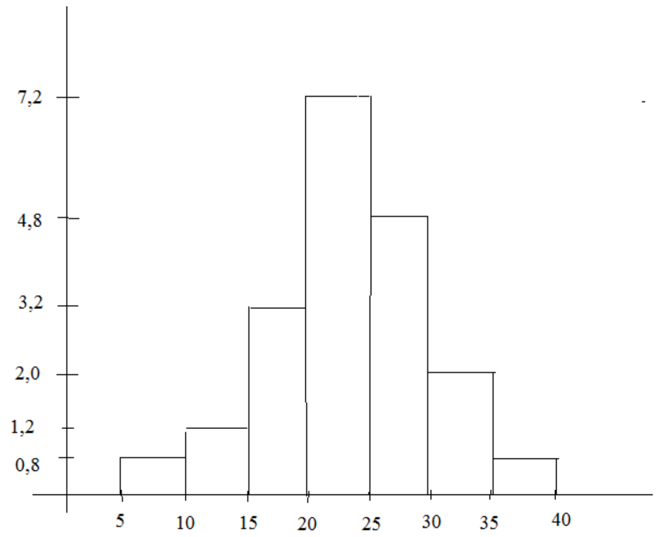


Рис. 47

Форма гістограми свідчить про те, що даний розподіл близький до нормального.

3.7. Точкові оцінки параметрів розподілів, вимоги до них

Нехай треба дослідити кількісну ознаку генеральної сукупності. Припустимо, що з деяких теоретичних міркувань вдалося встановити закон розподілу випадкової величини X. Тоді виникає задача оцінки параметрів, якими визначається цей розподіл. Наприклад, якщо X розподілена нормально, то такими параметрами можуть бути a та σ; якщо X розподілена за показниковим законом (див. п. 2.22), то параметром є λ, тощо.

Для оцінки цих параметрів дослідник має у своєму розпорядженні лише дані вибірки, які отримано в результаті спостережень. Якщо елементи вибірки $x_1, x_2, ..., x_n$ розглядати як незалежні випадкові величини, то оцінкою невідомого параметру буде функція від цих величин.

Означення. *Статистичною оцінкою невідомого параметра* випадкової величини X генеральної сукупності називається функція від випадкових величин $x_1, x_2, ..., x_n$ – елементів вибірки, що спостерігаються.

Щоб статистичні оцінки давали задовільні наближення невідомих параметрів, вони повинні задовольняти певним вимогам.

Нехай θ^* є статистичною оцінкою невідомого параметра θ теоретичного розподілу. Припустимо, що за вибіркою об'єма n знайдено оцінку θ_1^*. При інших вибірках того ж об'єму одержимо деякі інші оцінки $\theta_2^*,\ldots,\theta_k^*$, які, взагалі кажучи, відрізняються одна від одної. Тоді оцінку θ^* можна розглядати як випадкову величину, а числа $\theta_1^*,\theta_2^*,\ldots,\theta_k^*$ як її можливі значення.

Якщо числа $\theta_1^*,\theta_2^*,\ldots,\theta_k^*$ більше значення θ, то оцінка θ^* дає наближене значення θ з надлишком. У цьому випадку й математичне сподівання випадкової величини θ^* більше θ, тобто $M[\theta^*]>\theta$. Якщо θ^* дає оцінку θ з недостачею, то $M[\theta^*]<\theta$. Таким чином, використання статистичної оцінки, математичне сподівання якої не дорівнює параметру, який оцінюється, приводить до систематичних (одного знаку) похибок. Тому природно вимагати, щоб $M[\theta^*]=\theta$. Ця вимога застерігає систематичних похибок.

Означення. Статистичну оцінку θ^* параметра θ називають *незсунутою*, якщо $M[\theta^*]=\theta$.

Оцінку θ^* називають *зсунутою*, якщо $M[\theta^*]\neq\theta$.

Вимога незсунутості оцінки θ^* є недостатньою, адже можливі значення θ^* можуть бути значно розсіяні від свого середнього значення, тобто дисперсія $D[\theta^*]$ може бути великою. Тоді знайдена за даними однієї вибірки оцінка, наприклад θ_1^*, може значно відрізнятися від середнього значення θ^*, отже, і від параметра θ. Якщо вимагати, щоб дисперсія величини θ^* була малою, то можливість припуститися великої помилки буде виключена. Тому до статистичної оцінки виникає вимога про її ефективність.

Означення. Оцінка θ^* називається *ефективною*, якщо при заданому об'єму n вибірки має найменшу можливу дисперсію.

При розгляданні вибірок великого об'єму до статистичних оцінок пред'являють вимогу їх обґрунтованості.

Означення. Оцінка $\theta^*=\theta^*(n)$ невідомого параметра θ, яка отримана за вибіркою об'єма n, називається *обґрунтованою*, якщо $\forall\,\varepsilon>0$

$$\lim_{n\to\infty}P\left(\left|\theta^*(n)-\theta\right|<\varepsilon\right)=1.$$

Тобто при збільшенні об'єму вибірки подія, яка полягає в тому, що оцінка $\theta^*(n)$ буде мало відрізнятися від параметра θ, стає практично достовірною.

Якщо дисперсія незсунутої оцінки при $n\to\infty$ прямує до нуля, то така оцінка буде й обґрунтованою.

3.8. Вибіркове середнє

Одним з важливим засобів обробки даних є обчислення їх числових характеристик. Найбільш важливі з них: середнє значення або математичне сподівання, дисперсія, середньоквадратичне відхилення. Ці характеристики можуть бути обчислені наближено за даними вибірки. За аналогією з математичним сподіванням, дисперсією та середньоквадратичним відхиленням визначають вибіркові характеристики.

Нехай задано вибірку:

x_i	x_1	x_2	$\cdots$	x_k
n_i	n_1	n_2	$\cdots$	n_k

$n = n_1 + n_2 + \ldots + n_k$ – об'єм вибірки.

Означення. *Вибірковим середнім* називають величину

$$\overline{x}_B = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_k x_k}{n},$$

тобто середнє арифметичне елементів вибірки, причому кожен її елемент враховано стільки разів, яка його частота.

Вибіркове середнє є статистичним аналогом математичного сподівання і використовується в якості статистичної оцінки математичного сподівання. Нескладно довести, що ця оцінка є незсунутою. Дійсно, розглянемо для спрощення випадок, коли $n_1 = n_2 = ... = n_k = 1$, $k = n$. Будемо розглядати $x_1, x_2, ... x_n$ як незалежні однаково розподілені випадкові величини $X_1, X_2, ..., X_n$. Оскільки ці величини розподілені однаково, то вони мають однакові числові характеристики, зокрема, одне й те ж математичне сподівання, яке позначимо як a. Оскільки математичне сподівання середнього арифметичного однаково розподілених випадкових величин дорівнює математичному сподіванню кожної з величин (див. п. 2.9), то

$$M[\overline{x}_B] = M\left[\frac{X_1 + X_2 + ... + X_n}{n}\right] = a.$$

З огляду на те, що кожна з величин $X_1, X_2, ..., X_n$ має той самий розподіл, що й генеральна сукупність (яку ми також розглядаємо як випадкову величину), робимо висновок, що числові характеристики цих величин і генеральної сукупності однакові. Зокрема, математичне сподівання a кожної з величин дорівнює математичному сподіванню $M[X]$ генеральної сукупності, тобто:

$$M[X] = a.$$

Отже:

$$M[\overline{x}_B] = M[X],$$

тобто $\overline{x}_B$ є незсунутою оцінкою математичного сподівання.

Вибіркове середнє є також обґрунтованою оцінкою для математичного сподівання $M[X]$. Дійсно, припустивши, що величини $X_1, X_2, ..., X_n$ мають обмежені дисперсії, ми можемо до цих величин застосувати теорему Чебишова, а саме:

$$\lim_{n\to\infty} P\left(\left|\overline{x}_B - a\right| < \varepsilon\right) = 1 \quad \forall\varepsilon > 0.$$

Таким чином, якщо по декількох вибірках достатньо великого об'єму з однієї й тієї ж генеральної сукупності знайдено вибіркові середні, то вони будуть наближено дорівнювати одна одної. У цьому полягає властивість стійкості вибіркових середніх.

3.9. Вибіркова дисперсія

Знову розглянемо вибірку об'єму n:

x_i	x_1	x_2	$\cdots$	x_k
n_i	n_1	n_2	$\cdots$	n_k

$n = n_1 + n_2 + \ldots + n_k$.

Означення. *Вибірковою дисперсією* називається величина:

$$D_B = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - \overline{x}_B)^2.$$

Інакше кажучи це середнє арифметичне квадратів відхилень варіант від вибіркового середнього з урахуванням відповідних частот.

На практиці величину D_B, як правило, простіше обчислювати за формулою:

$$D_B = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - (\overline{x}_B)^2,$$

тобто середнє квадратів без квадрату середнього. Дійсно:

$$D_B = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - \overline{x}_B)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i^2 - 2\overline{x}_B x_i + (\overline{x}_B)^2) =$$

$$= \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \frac{2\overline{x}_B}{n}\sum_{i=1}^{k} n_i x_i + \frac{1}{n}(\overline{x}_B)^2 \sum_{i=1}^{k} n_i = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - 2\overline{x}_B \cdot \overline{x}_B + (\overline{x}_B)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - (\overline{x}_B)^2.$$

Крім вибіркової дисперсії користуються також вибірковим середньоквадратичним відхиленням.

Означення. *Вибірковим середньоквадратичним відхиленням* називається арифметичний квадратний корінь з вибіркової дисперсії:

$$\sigma_B = \sqrt{D_B}\,.$$

Вибіркова дисперсія є статистичним аналогом дисперсії випадкової величини і використовується в якості точкової оцінки дисперсії. Але можна довести, що ця оцінка є зсунутою, а саме:

$$M[D_B] = \frac{n-1}{n} D[X].$$

З огляду на це використовують також так звану *виправлену вибіркову дисперсію*:

$$s^2 = \frac{n}{n-1} D_B = \frac{1}{n-1}\sum_{i=1}^{k} n_i (x_i - \overline{x}_B)^2 .$$

При великих n дріб $n/(n-1)$ близький до 1, тому s^2 незначно відрізняється від D_B, хоча трохи більше її. В той же час:

$$M[s^2] = M\left[\frac{n}{n-1} D_B\right] = \frac{n}{n-1} M[D_B] = \frac{n}{n-1}\cdot\frac{n-1}{n} D[X] = D[X],$$

тобто s^2 є незсунутою оцінкою для дисперсії ознаки X генеральної сукупності.

Для оцінки середньоквадратичного відхилення ознаки X генеральної сукупності використовують σ_B, а також *виправлене вибіркове середньоквадратичне відхилення:*

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{k} n_i (x_i - \overline{x}_B)^2}\ .$$

Виправленими вибірковими дисперсією та середньоквадратичним відхиленням користуються, як правило, якщо $n < 30$.

3.10. Початкові та центральні емпіричні моменти

Означення. *Початковим емпіричним моментом* порядку m вибірки називають величину

$$\nu_m^* = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^m ,$$

тобто середнє арифметичне m-х степенів варіант вибірки.

Означення. *Центральним емпіричним моментом* порядку m називають величину:

$$\mu_m^* = \frac{1}{n}\sum_{i=1}^{k} n_i \left(x_i - \overline{x}_B \right)^m,$$

тобто середнє арифметичне m-х степенів відхилень варіант вибірки від вибіркового середнього.

Зокрема:

$$\nu_1^* = \overline{x}_B, \quad \mu_1^* = 0, \quad \mu_2^* = D_B.$$

Для центрального емпіричного моменту 2-го порядку маємо формулу:

$$\mu_2^* = \nu_2^* - \left(\nu_1^*\right)^2.$$

Для центральних емпіричних моментів 3-го та 4-го порядків нескладно довести справедливість формул:

$$\mu_3^* = \nu_3^* - 3\nu_1^*\nu_2^* + 2\left(\nu_1^*\right)^3,$$

$$\mu_4^* = \nu_4^* - 4\nu_3^*\nu_1^* + 6\nu_2^*\left(\nu_1^*\right)^2 - 3\left(\nu_1^*\right)^4,$$

які аналогічні відповідним формулам для центральних теоретичних моментів (див. п. 2.10).

Емпірічні початкові та центральні моменти є статистичними аналогами відповідно початкових та центральних теоретичних моментів випадкової величини. І вони використовуються в якості точкових оцінок теоретичних моментів.

3.11. Порівняльна таблиця характеристик випадкової величини та їх статистичних аналогів

Підсумовуючи викладене у пп. 3.1 – 3.10, можна скласти таблицю, до якої внесено основні характеристики випадкових величин та статистичні аналоги, що їм відповідають, або вибіркові характеристики.

Характеристики випадкових величин	*Вибіркові характеристики*
Закон розподілу випадкової величини	Статистичний розподіл вибірки
Значення випадкової величини	Варіанти вибірки
Ймовірності значень випадкової величини	Відносні частоти варіант вибірки
Функція розподілу випадкової величини	Емпірична функція розподілу
Щільність розподілу випадкової величини	Гістограма відносних частот вибірки

Математичне сподівання випадкової величини	Вибіркове середнє
Дисперсія випадкової величини	Вибіркова дисперсія
Середньоквадратичне відхилення випадкової величини	Вибіркове середньоквадратичне відхилення
Початкові та центральні теоретичні моменти	Початкові та центральні емпіричні моменти

3.12. Приклади на обчислення числових характеристик вибірки

Приклад 1. Задано вибірку:

x_i	2	5	7	10
n_i	16	12	8	14

Знайти вибіркове середнє $\overline{x}_B$, вибіркову дисперсію D_B, вибіркове середньоквадратичне відхилення σ_B, виправлену вибіркову дисперсію s^2, виправлене середньоквадратичне відхилення s.

Знайдемо об'єм вибірки: $n = 16 + 12 + 8 + 14 = 50$.

Далі знайдемо вибіркове середнє:

$$\overline{x}_B = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \frac{16\cdot 2 + 12\cdot 5 + 8\cdot 7 + 14\cdot 10}{50} = 5,76.$$

Вибіркову дисперсію знайдемо за формулою:

$$D_B = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \left(\overline{x}_B\right)^2 = \frac{16\cdot 4 + 12\cdot 25 + 8\cdot 49 + 14\cdot 100}{50} - 5,76^2 = 9,9424.$$

Тоді:

$$\sigma_B = \sqrt{D_B} = \sqrt{9,9424} \approx 3,153,$$

$$s^2 = \frac{n}{n-1} D_B = \frac{50}{49}\cdot 9,9424 \approx 10,145,$$

$$s \approx \sqrt{10,145} \simeq 3,185.$$

Розглянемо вибірку $x_1, x_2, ..., x_n$ та вибірку $u_1, u_2, ..., u_n$ того ж об'єму, причому $u_i = x_i - C$, де C – стала. Тоді очевидно, що

$$\overline{u}_B = \overline{x}_B - C, \quad D_B[u] = D_B[x].$$

Розглянемо тепер вибірку $v_1, v_2, \ldots, v_n$, причому $v_i = Cx_i$. Тоді:

$\overline{v}_B = C\,\overline{x}_B$, $\quad D_B[v] = C^2 D[x]$.

З огляду на ці співвідношення іноді для обчислення числових характеристик вибірки зручно перейти до так званих *умовних варіант*. Нехай, наприклад, $x_1, x_2, ..., x_n$ – великі числа, безпосередньо з якими складно вести обчислення. Тоді доцільно відняти від кожної варіанти одне й те ж число C, яке за порядком співпадає з варіантами, тобто перейти до умовних варіант $u_i = x_i - C$, та обчислити їх числові характеристики, зокрема, $\overline{u}_B$ та $D_B[u]$. А далі повернутися до числових характеристик початкових варіант за формулами:

$\overline{x}_B = \overline{u}_B + C$, $\quad D_B[x] = D_B[u]$.

<u>Приклад 2</u>. Знайти вибіркове середнє та вибіркову дисперсію для вибірки:

x_i	2560	2600	2620	2650	2700
n_i	2	3	10	4	1

Перейдемо до умовних варіант $u_i = x_i - 2620$:

u_i	-60	-20	0	30	80
n_i	2	3	10	4	1

Знайдемо:

$$\overline{u}_B = \frac{1}{n}\sum_{i=1}^{5} n_i u_i = \frac{-2\cdot 60 - 3\cdot 20 + 10\cdot 0 + 4\cdot 30 + 1\cdot 80}{20} = 1,$$

$$D_B[u] = \frac{1}{n}\sum_{i=1}^{5} n_i u_i^2 - \left(\overline{u}_B\right)^2 = \frac{2\cdot 3600 + 3\cdot 400 + 4\cdot 900 + 1\cdot 6400}{20} - 1 = 18399.$$

Тоді:

$\overline{x}_B = \overline{u}_B + 2620 = 2621$,

$D_B[x] = D_B[u] = 18399$.

Якщо варіанти $x_1, x_2, ..., x_n$ є десятковими дробами з l знаками після коми, то доцільно перейти до умовних варіант $v_i = 10^l x_i$. Тоді:

$$\overline{x}_B = \frac{1}{10^l}\overline{v}_B, \quad D_B[x] = \frac{1}{10^{2l}} D_B[v].$$

<u>Приклад 3</u>. Задано вибірку:

x_i	0,01	0,05	0,09
n_i	2	3	5

Знайти $\overline{x}_B$, $D_B[x]$.

-*

Перейдемо до умовних варіант $v_i = 100 x_i$:

v_i	1	5	9
n_i	2	3	5

Знайдемо:

$$\overline{v}_B = \frac{2\cdot 1 + 3\cdot 5 + 5\cdot 9}{10} = 6,2,$$

$$D_B[v] = \frac{2\cdot 1 + 3\cdot 25 + 5\cdot 81}{10} - 6,2^2 = 9,76.$$

Тоді:

$$\overline{x}_B = \frac{1}{100}\overline{v}_B = 0,062,$$

$$D_B[x] = \frac{1}{10000}D_B[v] = 0,000976.$$

Часто доцільно комбінувати ці два способи.

<u>Приклад 4</u>. Задано вибірку:

x_i	23,5	26,1	28,2	30,4
n_i	2	3	4	1

Знайти $\overline{x}_B$, $D_B[x]$.

Перейдемо до умовних варіант $v_i = 10x_i$:

v_i	235	261	282	304
n_i	2	3	4	1

А тепер до варіант $u_i = v_i - 268$:

u_i	-33	-7	14	36
n_i	2	3	4	1

Знайдемо:

$$\overline{u}_B = \frac{-2\cdot 33 - 3\cdot 7 + 4\cdot 14 + 1\cdot 36}{10} = 0,5,$$

$$D_B[u] = \frac{2\cdot 33^2 + 3\cdot 7^2 + 4\cdot 14^2 + 1\cdot 36^2}{10} - 0,5^2 = 440,25.$$

Тоді:

$$\overline{v}_B = \overline{u}_B + 268 = 268,5,$$

$$D_B[v] = D_B[u] = 440,25,$$

$$\overline{x}_B = \frac{1}{10}\overline{v}_B = 26,85,$$

$$D_B[x] = \frac{1}{100} D_B[v] = 0{,}01 \cdot 440{,}25 = 4{,}4025.$$

3.13. Метод моментів точкової оцінки параметрів розподілу

Можна довести, що початкові та центральні емпіричні моменти є обґрунтованими оцінками відповідно початкових та центральних теоретичних моментів того ж порядку. На цьому ґрунтується метод моментів точкової оцінки параметрів розподілу, запропонований К. Пірсоном. Він полягає в тому, що теоретичні моменти дорівнюються відповідним теоретичним моментам.

Нехай, наприклад, теоретичний розподіл визначається одним параметром θ, тобто щільність розподілу має вигляд $f(x, \theta)$. Треба знайти точкову оцінку для параметра θ. Для цього дорівняємо початковий теоретичний момент 1-го порядку ознаки X генеральної сукупності початковому емпіричному моменту 1-го порядку, тобто:

$$\nu_1[X] = \nu_1^*.$$

Враховуючи, що $\nu_1[X] = M[X]$, $\nu_1^* = \overline{x}_B$, отримаємо:

$$M[X] = \overline{x}_B. \qquad (*)$$

Для неперервної випадкової величини математичне сподівання знаходиться за формулою:

$$M[X] = \int_{-\infty}^{+\infty} x f(x, \theta)\, dx = \varphi(\theta),$$

тобто це математичне сподівання є функцією параметра θ, тому рівність (*) можна розглядати як рівняння відносно цього параметра:

$$\varphi(\theta) = \overline{x}_B.$$

Розв'язавши це рівняння відносно θ, знайдемо його точкову оцінку θ^*, яка є функцією від вибіркового середнього, а отже, й варіант вибірки:

$$\theta^* = \psi(x_1, x_2, ..., x_n).$$

Приклад. Задано вибірку:

x_i	x_1	x_2	$\cdots$	x_k
n_i	n_1	n_2	$\cdots$	n_k

$n = n_1 + n_2 + ... + n_k$ – об'єм вибірки. За цією вибіркою методом моментів знайти точкову оцінку невідомого параметра λ показникового розподілу, який має щільність

$$f(x,\lambda) = \lambda e^{-\lambda x} \quad (x \geq 0).$$

Враховуючи те, що математичне сподівання показникового розподілу дорівнює $1/\lambda$ (див. п. 2.23), дорівняємо математичне сподівання до вибіркового середнього (тобто дорівняємо теоретичний початковий момент 1-го порядку емпіричному початковому моменту 1-го порядку):

$$\frac{1}{\lambda} = \overline{x}_B = \frac{1}{n}\sum_{i=1}^{k} n_i x_i .$$

Звідси:

$$\lambda = \frac{1}{\overline{x}_B}.$$

Таким чином, точкова оцінка параметра λ:

$$\lambda^* = \frac{1}{\overline{x}_B} = \frac{n}{\sum\limits_{i=1}^{k} n_i x_i}.$$

Нехай тепер розподіл визначається двома параметрами θ_1, θ_2, тобто щільність розподілу має вигляд $f(x, \theta_1, \theta_2)$. Тоді необхідно мати два рівняння для знаходження цих параметрів. Дорівняємо математичне сподівання вибірковому середньому, а дисперсію – вибірковій дисперсії (тобто теоретичний початковий момент 1-го порядку емпіричному початковому моменту 1-го порядку, а теоретичний центральний момент 2-го порядку – емпіричному центральному моменту 2-го порядку):

$$M[X] = \overline{x}_B, \; D[X] = D_B.$$

Наприклад, для параметрів a, σ нормально розподіленої випадкової величини матимемо наступні точкові оцінки:

$$a^* = \overline{x}_B, \; \sigma^* = \sqrt{D_B} .$$

3.14. Метод найбільшої правдоподібності

Нехай X – ДВВ, яка внаслідок n випробувань прийняла значення $x_1, x_2, ..., x_n$ (для спрощення вважатимемо, що всі частоти дорівнюють 1. Припустимо, що задано закон розподілу величини X, але невідомо параметр θ, яким цей розподіл визначається. Треба знайти його точкову оцінку.

Позначимо як $p(x_i, \theta)$ ймовірність того, що внаслідок випробування величина X прийме значення x_i $(i = 1, 2, ..., n)$.

Означення. *Функцією правдоподібності* ДВВ X називається функція аргумента θ:

$$L(x_1, x_2, ..., x_n; \theta) = p(x_1, \theta) p(x_2, \theta) \cdots p(x_n, \theta),$$

де $x_1, x_2, ..., x_n$ – фіксовані числа.

В якості точкової оцінки параметра θ приймають таке значення $\theta^* = \theta^*(x_1, x_2, ..., x_n)$, за яким функція правдоподібності досягає максимуму. Справді, якщо функція L досягає максимуму, то у певному розумінні це буде означати й досягання максимумів ймовірностей $p(x_1,\theta), p(x_2,\theta), ..., p(x_n,\theta)$, а це фактично означає, що при даному значенні параметра $\theta = \theta^*$ теоретичні очікування близькі до емпіричних даних.

Оцінку θ^* називають *оцінкою найбільшої правдоподібності*.

Іноді замість максимуму функції L простіше шукати максимум функції $\ln L$ – ці дві функції досягають максимуму при одному й тому ж значенні θ. Функцію $\ln L$ називають *логарифмічною функцією правдоподібності*. Точку максимуму функції L можна шукати так:

1) знайти похідну $d \ln L / d\theta$;

2) дорівняти цю похідну до нуля та знайти критичну точку θ^* з рівняння (рівняння правдоподібності):

$$\frac{d \ln L}{d\theta} = 0;$$

3) знайти похідну 2-го порядку $d^2 \ln L / d\theta^2$; якщо при $\theta = \theta^*$ ця похідна від'ємна, то θ^* – точка максимуму.

Ця точка максимуму θ^* приймається в якості точкової оцінки параметра θ.

Приклад. Знайти методом найбільшої правдоподібності оцінку параметра λ розподілу Пуассона:

$$P_m(X = x_j) = \frac{\lambda^{x_j} e^{-\lambda}}{x_j!},$$

де m – число проведених випробувань; x_j – число появ події в j-му $(j = 1, 2, ..., n)$ досліді (кожний дослід складається з m випробувань).

Складемо функцію правдоподібності (тут $\theta = \lambda$):

$$L = p(x_1,\lambda) p(x_2,\lambda) \cdots p(x_n,\lambda) = \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} \cdots \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} = \frac{\lambda^{x_1 + x_2 + ... + x_n}}{x_1! x_2! \cdots x_n!} e^{-n\lambda}.$$

Знайдемо логарифмічну функцію правдоподібності:

$$\ln L = (x_1 + x_2 + ... + x_n) \ln \lambda - n\lambda - \ln(x_1! x_2! \cdots x_n!).$$

Знайдемо її похідну:

$$\frac{d\ln L}{d\lambda}=\frac{x_1+x_2+...+x_n}{\lambda}-n.$$

Складемо рівняння правдоподібності:

$$\frac{x_1+x_2+...+x_n}{\lambda}-n=0,$$

з якого отримуємо критичну точку:

$$\lambda^*=\frac{x_1+x_2+...+x_n}{n}=\overline{x}_B.$$

Знайдемо похідну 2-го порядку функції $\ln L$:

$$\frac{d^2\ln L}{d\lambda^2}=-\frac{x_1+x_2+...+x_n}{\lambda^2}.$$

Очевидно, що ця похідна від'ємна, отже $\lambda^*=\overline{x}_B$ – точка максимуму. Таким чином, в якості точкової оцінки λ^* параметра λ приймаємо вибіркове середнє:

$$\lambda^*=\overline{x}_B.$$

Нехай тепер X – НВВ, яка внаслідок випробування прийняла значення $x_1,x_2,...,x_n$. Припустимо, що вигляд щільності розподілу $f(x,\theta)$ відомо, але невідомо параметр θ, яким ця функція визначається.

Означення. Функцією правдоподібності НВВ X називається функція аргументу θ:

$$L(x_1,x_2,...,x_n;\theta)=f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta),$$

де $x_1,x_2,...,x_n$ – фіксовані числа.

Оцінку найбільшої правдоподібності шукають так саме, як у випадку ДВВ.

Якщо щільність розподілу визначається двома параметрами θ_1,θ_2, то функція правдоподібності є функцією двох аргументів θ_1,θ_2:

$$L=f(x_1;\theta_1,\theta_2)f(x_2;\theta_1,\theta_2)\cdots f(x_n;\theta_1,\theta_2).$$

Для знаходження $\max L$ розв'язують систему рівнянь:

$$\frac{\partial\ln L}{\partial\theta_1}=0,\ \frac{\partial\ln L}{\partial\theta_2}=0,$$

а потім користуються достатніми умовами максимуму функції двох змінних.

Приклад. Методом найбільшої правдоподібності знайти точкові оцінки параметрів a,σ нормального розподілу:

$$f(x,a,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}},$$

якщо внаслідок n випробувань величина X прийняла значення $x_1,x_2,...,x_n$.

Складемо функцію правдоподібності:

$$L=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_1-a)^2}{2\sigma^2}}\cdot\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_2-a)^2}{2\sigma^2}}\cdots\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_n-a)^2}{2\sigma^2}}=$$

$$=\frac{1}{\left(\sigma\sqrt{2\pi}\right)^n}\exp\left(-\frac{(x_1-a)^2+(x_2-a)^2+\ldots+(x_n-a)^2}{2\sigma}\right).$$

Знайдемо логарифмічну функцію правдоподібності:

$$\ln L=-n\ln\sigma-n\ln\sqrt{2\pi}-\frac{(x_1-a)^2+(x_2-a)^2+\ldots+(x_n-a)^2}{2\sigma^2}.$$

Знайдемо частині похідні за змінними a і σ:

$$\frac{\partial\ln L}{\partial a}=\frac{2(x_1-a)+2(x_2-a)+\ldots+2(x_n-a)}{2\sigma^2}=\frac{x_1+x_2+\ldots+x_n-na}{\sigma^2};$$

$$\frac{\partial\ln L}{\partial\sigma}=-\frac{n}{\sigma}+\frac{(x_1-a)^2+(x_2-a)^2+\ldots+(x_n-a)^2}{\sigma^3}.$$

Складемо систему:

$$\frac{\partial\ln L}{\partial a}=0,\ \frac{\partial\ln L}{\partial\sigma}=0.$$

У даному випадку вона має вигляд:

$$\frac{x_1+x_2+\ldots+x_n-na}{\sigma^2}=0,$$

$$-\frac{n}{\sigma}+\frac{(x_1-a)^2+(x_2-a)^2+\ldots+(x_n-a)^2}{\sigma^3}=0.$$

З неї отримаємо:

$$a^*=\frac{x_1+x_2+\ldots+x_n}{n}=\overline{x}_B,$$

$$\sigma^*=\sqrt{\frac{(x_1-a)^2+(x_2-a)^2+\ldots+(x_n-a)^2}{n}}=\sqrt{D_B}=\sigma_B.$$

Це й є оцінки найбільшої правдоподібності. Як бачимо, вони співпадають зі статистичними аналогами відповідно математичного сподівання та середньоквадратичного відхилення.

3.15. Інші вибіркові характеристики

Крім вибіркового середнього та вибіркової дисперсії використовуються інші характеристики вибірки.

Означення. *Модою* M_0 називається варіанта, яка має найбільшу частоту.

Для вибірки

x_i	1	4	7	9
n_i	5	1	20	6

мода дорівнює 7.

Означення. *Медіаною* m_e називається число, яке ділить статистичний розподіл вибірки на дві частини, які співпадають за кількістю варіант.

Наприклад, для вибірки

x_i	2	3	5	6	7
n_i	1	2	1	1	1

медіана дорівнює: $m_e = (3+5)/2 = 4$.

Означення. *Розмахом варіювання* називається різниця між найбільшою та найменшою варіантами:

$$R = x_{\max} - x_{\min}.$$

Для попередньої вибірки: $R = 7 - 2 = 5$.

Означення. *Середнім абсолютним відхиленням* називається величина

$$\sigma_{\text{abs}} = \frac{1}{n}\sum_{i=1}^{k} n_i \,|\, x_i - \overline{x}_B \,|.$$

Величина σ_{abs} використовується для характеристики розсіяння статистичного розподілу вибірки.

Рівностями

$$a^* = \frac{\overset{*}{\mu}_3}{\sigma_B^3}, \quad e^* = \frac{\overset{*}{\mu}_4}{\sigma_B^4} - 3,$$

визначаються відповідно *асиметрія* та *ексцес* емпіричного розподілу.

3.16. Інтервальні оцінки параметрів розподілів
Довірчий інтервал і довірча ймовірність

Точкова оцінка параметрів розподілу визначається одним числом. Точкові оцінки є випадковими величинами, їх можна вважати первинними результатами обробки вибірки, адже невідомо, з якою точністю кожна з них оцінює відповідну числову характеристику генеральної сукупності. Якщо об'єм вибірки досить великий, то точкові оцінки задовольняють практичні потреби точності. Але якщо об'єм вибірки малий, то точкові оцінки можуть давати значні похибки. Тому при малих об'ємах вибірки користуються інтервальними оцінками.

Означення. *Інтервальною* називається оцінка, яка визначається двома числами – кінцями інтервалу.

Інтервальні оцінки дозволяють встановити точність та надійність оцінок. Розглянемо ці поняття детальніше.

Нехай знайдена за даними вибірки статистична оцінка θ^* буде оцінкою невідомого параметра θ. Зрозуміло, що θ^* тим точніше визначає θ, чим менше величина $|\theta^* - \theta|$. Іншими словами, якщо $\delta > 0$ і $|\theta^* - \theta| < \delta$, то чим менше δ, тим більш точна оцінка. Тому число δ характеризує точність оцінки.

Але статистичні методи не дозволяють категорично стверджувати, що оцінка θ^* задовольняє нерівність $|\theta^* - \theta| < \delta$, можна лише казати про ймовірність γ, з якою ця нерівність виконується.

Означення. *Довірчою ймовірністю* (або *надійністю*) статистичної оцінки θ^* параметра θ називається ймовірність γ, з якою виконано нерівність $|\theta^* - \theta| < \delta$.

Найчастіше число γ задається наперед, причому в якості γ береться число, близьке до 1, наприклад 0,95; 0,99; 0,999. У задачах геології зазвичай використовується значення $\gamma = 0,95$.

Отже:

$$P\left(|\theta^* - \theta| < \delta\right) = \gamma. \qquad (*)$$

Тепер задача полягає в тому, щоб за даним γ з нерівності (*) визначити число δ. Тим самим визначиться інтервал, якому з ймовірністю γ належить невідомий параметр θ.

Замінимо нерівність $|\theta^* - \theta| < \delta$ рівносильною нерівністю:

$$\theta^* - \delta < \theta < \theta^* + \delta.$$

Тоді:

$$P\left(\theta^* - \delta < \theta < \theta^* + \delta\right) = \gamma.$$

Таким чином, той інтервал, якому з ймовірністю γ належить невідомий параметр θ, є $\left(\theta^* - \delta, \theta^* + \delta\right)$.

Означення. Інтервал $\left(\theta^* - \delta, \theta^* + \delta\right)$ називається *довірчим* для невідомого параметра θ, якщо він покриває параметр θ із заданою надійністю γ.

Зауваження. Кінці довірчого інтервалу є випадковими величинами.

3.17. Довірчий інтервал для оцінки математичного сподівання нормального розподілу

Нехай кількісна ознака X генеральної сукупності розподілена за нормальним законом, причому середньоквадратичне відхилення σ цього розподілу відомо. Треба знайти довірчий інтервал, що покриває математичне сподівання a генеральної сукупності із заданою надійністю γ.

Згідно із властивістю нормально розподіленої випадкової величини X (див. п. 2.27) маємо:

$$P\left(|X - a| < \delta\right) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

Можна довести, що, якщо величина X розподілена нормально, то вибіркове середнє $\overline{x}_B$, також розподілено нормально, причому (див. п. 2.9):

$$M[\overline{x}_B] = a, \quad \sigma[\overline{x}_B] = \frac{\sigma}{\sqrt{n}},$$

де n – об’єм вибірки.

Тому

$$P\left(|\overline{x}_B - a| < \delta\right) = 2\Phi\left(\frac{\delta}{\sigma/\sqrt{n}}\right) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right) = 2\Phi(t),$$

де $t = \delta\sqrt{n}/\sigma$. Тоді $\delta = t\sigma/\sqrt{n}$, і можемо написати:

$$P\left(|\overline{x}_B - a| < \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma.$$

Або:

$$P\left(\overline{x}_B - \frac{t\sigma}{\sqrt{n}} < a < \overline{x}_B + \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma.$$

Таким чином, з надійністю γ можна стверджувати, що довірчий інтервал

$$\left(\overline{x}_B - \frac{t\sigma}{\sqrt{n}}, \overline{x}_B + \frac{t\sigma}{\sqrt{n}}\right) \qquad (*)$$

покриває невідомий параметр a. Точність оцінки:

$$\delta = \frac{t\sigma}{\sqrt{n}}. \qquad (**)$$

Число t визначається рівністю $2\Phi(t) = \gamma$, або

$$\Phi(t) = \frac{\gamma}{2}.$$

За відомим числом γ за таблицею інтегральної функції Лапласа (див. Додаток, Таблиця 2) знаходять число t, якому відповідає значення $\gamma/2$ функції Лапласа. Наприклад, якщо $\gamma = 0,95$, то $\gamma/2 = 0,475$, і за таблицею знаходимо: $t = 1,96$.

Зауваження. З формули (**) випливає, що при зростанні об'єму вибірки n число δ зменшується, а це означає, що точність оцінки збільшується. Коли збільшується надійність γ, внаслідок властивості зростання функції Φ, зростає t, отже зростає δ, а це означає, що точність оцінки зменшується.

Приклад. Вибірка з великої партії зразків мінералів містить 100 зразків. Середній вміст деякого хімічного елементу в цих зразках з'явився 13%. Знайти з надійністю 0,95 довірчий інтервал для середнього вмісту a хімічного елементу в зразках всієї партії, якщо відомо, що середньоквадратичне відхилення вмісту елементу $\sigma = 2\%$.

Маємо $\overline{x}_B = 13$, $\sigma = 2$, $\gamma = 0,95$, отже $t = 1,96$, $n = 100$. Тому, згідно з формулою (*), довірчий інтервал такий:

$$\left(13 - \frac{1,96 \cdot 2}{\sqrt{100}}, 13 + \frac{1,96 \cdot 2}{\sqrt{100}}\right),$$

тобто

$$(12,608; 13,392).$$

Якщо середньоквадратичне відхилення σ генеральної сукупності невідомо, то для знаходження довірчого інтервалу користуються формулою:

$$\left(\overline{x}_B - \frac{t_\gamma s}{\sqrt{n}}, \overline{x}_B + \frac{t_\gamma s}{\sqrt{n}}\right), \qquad (***)$$

де s – виправлене вибіркове середньоквадратичне відхилення, а t_γ – спеціальний параметр, який знаходиться за таблицею (див. Додаток, Таблиця 3) за відомими γ і n. Наприклад, для $\gamma = 0,95$, $n = 50$: $t_\gamma = 2,009$.

Для оцінки середньоквадратичного відхилення σ нормального розподілу використовується наступний довірчий інтервал:

$$(s(1-q),\, s(1+q)), \qquad (****)$$

де s – виправлене вибіркове середньоквадратичне відхилення, а q – параметр, який знаходиться за таблицею (див. Додаток, Таблиця 4) за відомими γ і n. Наприклад, для $\gamma = 0{,}95$, $n = 50$: $q = 0{,}21$.

3.18. Статистична перевірка статистичних гіпотез

У розглянутих в попередніх параграфах задачах ми оцінювали параметри розподілів, вважаючи при цьому, що вигляд самого розподілу відомий. Наприклад, при отриманні інтервальної оцінки у попередньому параграфі ми припускали, що величина розподілена за нормальним законом. Разом з цим в багатьох задачах закон розподілу випадкової величини є невідомим. У зв'язку з цим доводиться будувати різні припущення (гіпотези) щодо вигляду цього закону. Припустимо, що ми маємо певні підстави вважати, що функція розподілу дорівнює $F(x)$, де $F(x)$ – деяка відома функція. Таке припущення називається *статистичною гіпотезою*. Щоб перевірити достовірність цієї гіпотези, треба порівняти як узгоджуються експериментальні дані (тобто елементи вибірки) з гіпотетичною функцією розподілу $F(x)$. Ми знаємо, що статистичним аналогом функції розподілу є емпірична функція розподілу $F_n(x)$ (див. п. 3.5). Тому шукану відповідність можна оцінити як близькість емпіричної функції розподілу $F_n(x)$ та гіпотетичної функції розподілу $F(x)$. У зв'язку з цим вводиться деяка величина d, яка й характеризує цю близькість. Величину d можна обирати різними способами, у відповідності з якими отримуються різні критерії для перевірки гіпотези, що нас цікавить. Такі критерії називаються *критеріями узгодженості*. Наприклад, можна покласти:

$$d = \sup_x |F_n(x) - F(x)|.$$

Цей критерій називається *критерієм Колмогорова*. Або можна покласти:

$$d = \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 g(x)dx,$$

де

$$g(x) \ge 0, \quad \int_{-\infty}^{+\infty} g(x)dx < +\infty.$$

Цей критерій називається *критерієм Мізеса*.

Чим менше величина d, тим ближче $F_n(x)$ до $F(x)$, тим краще узгоджується статистична гіпотеза з даними вибірки.

На практиці критеріями узгодженості користуються так. Задається деяке число $\alpha > 0$ настільки мале, щоб подію, яка відбувається із ймовірністю α, можна вважати практично неможливою. За цим α вибирається число d_0 з наступної умови: $P(d > d_0) = \alpha$. Число α називається *рівнем значущості*. Число d_0 називається *критичною точкою*. За даними вибірки будуємо функцію $F_n(x)$ і знаходимо величину d відповідно прийнятому критерію узгодженості. Порівнюємо величину d з критичною точкою d_0. Якщо вийде, що $d > d_0$, то це означає, що припущення про справедливість гіпотези привело до висновку, що відбулася практично неможлива подія, і, таким чином, гіпотеза спростується експериментом. Тоді її слід відкинути. Якщо $d \le d_0$, то гіпотеза не суперечить даним вибірки, підстав для її відкидання нема, і вона може бути прийнята.

Слід відмітити, що спростування гіпотези при $d > d_0$ не означає її логічного спростування, а також підтвердження гіпотези у випадку $d \le d_0$ не означає логічного доведення справедливості гіпотези. Дійсно, подія $d > d_0$ може відбутися і у випадку справедливості гіпотези, але оскільки α мале, то на практиці цією можливістю можна нехтувати. Також і подія $d \le d_0$ може відбутися й тоді, коли гіпотеза невірна, тому її необхідно перевірити за допомогою інших критеріїв.

Очевидно, що величина d є випадковою (вона залежить від даних вибірки), та її розподіл залежить від об'єму вибірки n. Тому знаходження цього розподілу при фіксованих значеннях n достатньо важке та недоцільне. Замість цього обчислюють граничний (при $n \to \infty$) розподіл величини d та використовують його в якості наближеного значення для розподілу величини d при фіксованих, але досить великих значеннях n.

3.19. Критерій Пірсона перевірки статистичної гіпотези про нормальний розподіл випадкової величини

При отриманні критерію для перевірки гіпотези, яка полягає в тому, що функцією розподілу випадкової величини є певна функція $F(x)$, ми вводимо міру d відхилення емпіричної функції розподілу $F_n(x)$ від гіпотетичної функції розподілу $F(x)$. В якості такої міри часто використовується міра, яку введено К. Пірсоном.

Нехай задано вибірку об'єма n. Оберемо деякий інтервал (a,b), у якому містяться всі елементи вибірки. Розіб'ємо цей інтервал на декілька частин (для спрощення вважатимемо їх рівними). Позначимо $s_0 = a$, $s_r = b$ (r – число интервалів). Частинні проміжки позначатимемо як $[s_0, s_1]$, $[s_1, s_2], \ldots$, $[s_i, s_{i+1}], \ldots$, $[s_{r-1}, s_r]$.

Припустимо, що випадкову величину X, що досліджується, розподілено нормально із параметрами a і σ. Тоді (див. п. 2.26):

$$p_i = P(s_i < X < s_{i+1}) = \Phi\left(\frac{s_{i+1} - a}{\sigma}\right) - \Phi\left(\frac{s_i - a}{\sigma}\right).$$

Оскільки параметри a та σ завчасно невідомі, то їх наближено можна замінити їх статистичними аналогами, відповідно $\overline{x}_B$ та σ_B. Тоді:

$$p_i \approx \Phi\left(\frac{s_{i+1} - \overline{x}_B}{\sigma_B}\right) - \Phi\left(\frac{s_i - \overline{x}_B}{\sigma_B}\right),\ i = 0, 1, \ldots, r-1.$$

Знаючи ймовірності p_i, знайдемо для кожного частинного проміжку величину np_i. Величини np_i називаються *теоретичними частотами* – вони показують, скільки елементів вибірки, скоріше всього, потрапило б до відповідного інтервалу, якби випадкова величина X дійсно була б розподілена нормально з параметрами a та σ.

Далі для кожного частинного проміжку знайдемо величину ν_i, яка показує, скільки дійсно елементів вибірки потрапило до цього частинного проміжку. Величини ν_i називаються *емпіричними частотами*. Якщо гіпотеза про нормальний розподіл правдоподібна, то емпіричні частоти не повинні значно відрізнятися від теоретичних частот, тобто має бути: $\nu_i \approx np_i$. Тому треба ввести до розгляду величину, яка характеризувала б ступінь близькості цих частот. За пропозицією К. Пірсона такою мірою є:

$$\chi^2 = \sum_{i=0}^{r-1} \frac{(\nu_i - np_i)^2}{np_i}\ .$$

Це й є міра d відхилення емпіричної функції розподілу від теоретичної. Користуються цією мірою наступним чином. Визначається так зване число ступенів вільності $k = r - 3$ (r – кількість частинних проміжків). Задається рівень значущості α, і в залежності від α і k за таблицею 5 (див. Додаток) визначають критичну точку $\chi^2_{kp} = \chi^2_{kp}(\alpha, k)$. Якщо знайдена за даними вибірки величина χ^2 більше, ніж χ^2_{kp}, то це буде означати, що розбіжність між емпіричними та теоретичними частотами суттєва, і у цьому випадку гіпотеза про нормальний розподіл має бути відкинута. Якщо $\chi^2 \le \chi^2_{kp}$, то це означає, що суттєвої розбіжності між емпіричними та теоретичними частотами нема, і тому гіпотезу може бути прийнято.

3.20. Приклад перевірки статистичної гіпотези про нормальний розподіл випадкової величини за допомогою критерію Пірсона

Нехай за допомогою критерію Пірсона на рівні значущості $\alpha = 0,01$ треба перевірити гіпотезу про нормальний розподіл випадкової величини за даними наступної вибірки об'єму $n = 50$:

0,60	1,64	2,07	2,20	2,31	2,81	2,82	2,91	3,41	3,42
3,54	3,77	3,90	4,23	4,33	4,34	4,57	5,08	5,15	5,49
5,50	5,51	5,64	5,66	5,77	5,96	6,08	6,09	6,13	6,16
6,31	6,38	6,39	6,59	6,68	7,18	7,41	7,44	7,72	8,12
8,29	8,42	8,55	8,87	8,93	9,24	9,52	9,56	10,07	10,72

Знайдемо числові характеристики цієї вибірки:

$\overline{x}_B = 5,7896$, $D_B = 5,6886$, $\sigma_B = 2,3851$, $s^2 = 5,8047$, $s = 2,4093$.

У припущенні, що випадкова величина розподілена нормально, знайдемо з надійністю $\gamma = 0,95$ довірчі інтервали для оцінки математичного сподівання та середньоквадратичного відхилення випадкової величини. Довірчий інтервал для оцінки математичного сподівання знаходимо за формулою (***) п. 3.17, де значення параметра $t_\gamma = t_\gamma(\gamma, n)$ для $\gamma = 0,95$, $n = 50$ дорівнює: $t_\gamma = 2,009$. Підставивши до цієї формули числові дані, отримуємо:

$(5,105; 6,4741)$.

Довірчий інтервал для оцінки середньоквадратичного відхилення знаходимо за формулою (****) п. 3.17, де значення параметра $q = q(\gamma, n)$ для $\gamma = 0,95$, $n = 50$ дорівнює: $q = 0,21$. Підставивши числові дані, отримуємо:

(1,90; 2,92).

Перейдемо тепер до перевірки статистичної гіпотези про нормальний розподіл. Оберемо відрізок $[0,12]$, до якого входять всі елементи вибірки, і розіб'ємо його на 6 частинних проміжків довжиною $h = 2$: $[0,2]$, $[2,4]$, $[4,6]$, $[6,8]$, $[8,10]$, $[10,12]$. Для кожного проміжку знайдемо ймовірність, що випадкова величина X прийме значення з цього проміжку, за формулою:

$$p_i \approx \Phi\left(\frac{s_{i+1} - \overline{x}_B}{\sigma_B}\right) - \Phi\left(\frac{s_i - \overline{x}_B}{\sigma_B}\right),\ i = 0,1,2,3,4,5.$$

Маємо:

$$p_1 \approx \Phi\left(\frac{2-5,7896}{2,3851}\right)-\Phi\left(\frac{0-5,7896}{2,3851}\right) \approx \Phi(-1,59)-\Phi(-2,43)=$$

$$=\Phi(2,43)-\Phi(1,59)=0,4924-0,4441=0,0483,$$

$$p_2 \approx \Phi\left(\frac{4-5,7896}{2,3851}\right)-\Phi\left(\frac{2-5,7896}{2,3851}\right) \approx \Phi(-0,75)-\Phi(-1,59)=$$

$$=\Phi(1,59)-\Phi(0,75)=0,4441-0,2734=0,1707,$$

$$p_3 \approx \Phi\left(\frac{6-5,7896}{2,3851}\right)-\Phi\left(\frac{4-5,7896}{2,3851}\right) \approx \Phi(0,09)-\Phi(-0,75)=$$

$$=\Phi(0,09)+\Phi(0,75)=0,0359+0,2734=0,3093,$$

$$p_4 \approx \Phi\left(\frac{8-5,7896}{2,3851}\right)-\Phi\left(\frac{6-5,7896}{2,3851}\right) \approx \Phi(0,93)-\Phi(0,09)=$$

$$=0,3238-0,0359=0,2879,$$

$$p_5 \approx \Phi\left(\frac{10-5,7896}{2,3851}\right)-\Phi\left(\frac{8-5,7896}{2,3851}\right) \approx \Phi(1,77)-\Phi(0,93)=$$

$$=0,4616-0,3238=0,1378,$$

$$p_6 \approx \Phi\left(\frac{12-5,7896}{2,3851}\right)-\Phi\left(\frac{10-5,7896}{2,3851}\right) \approx \Phi(2,60)-\Phi(1,77)=$$

$$=0,4953-0,4616=0,0337.$$

Тепер для кожного частинного проміжку знайдемо теоретичні частоти np_i:

$np_1=50\cdot 0,0483=2,42$,

$np_2=50\cdot 0,1707=8,53$,

$np_3=50\cdot 0,3093=15,46$,

$np_4=50\cdot 0,2879=14,39$,

$np_5=50\cdot 0,1378=6,89$,

$np_6=50\cdot 0,0337=1,68$.

Далі для кожного частинного проміжку знайдемо емпіричні частоти ν_i, тобто підрахуємо, скільки елементів вибірки потрапило до цього частинного проміжку. Отримаємо:

$\nu_1=2$, $\nu_2=11$, $\nu_3=13$, $\nu_4=13$, $\nu_5=9$, $\nu_6=2$.

Зведемо отримані дані до наступної таблиці:

Частинні проміжки	Емпірічні частоти ν_i	Ймовірності p_i	Теоретичні частоти np_i	$(\nu_i - np_i)^2$	$\dfrac{(\nu_i - np_i)^2}{np_i}$
$[0,2]$	2	0,0483	2,42	0,1764	0,073
$[2,4]$	11	0,1707	8,53	6,1	0,715
$[4,6]$	13	0,3093	15,46	6,05	0,391
$[6,8]$	13	0,2879	14,39	1,93	0,134
$[8,10]$	9	0,1378	6,89	4,45	0,646
$[10,12]$	2	0,0337	1,68	0,01	0,06

Обчислимо суму елементів останнього стовпця таблиці, тобто

$$\chi^2 = \sum_{i=0}^{5} \frac{(\nu_i - np_i)^2}{np_i} = 2,019 .$$

Порівняємо це значення з критичною точкою χ^2_{kp}, яка знаходиться за таблицею (див. Додаток, Таблиця 5) для числа ступенів вільності $k = 6 - 3 = 3$ та рівня значущості $\alpha = 0,01$: $\chi^2_{kp} = 11,3$. Оскільки $\chi^2 < \chi^2_{kp}$, то гіпотеза про нормальний розподіл випадкової величини приймається.

За першими двома стовпцями наведеної таблиці нескладно побудувати гістограму частот вибірки:

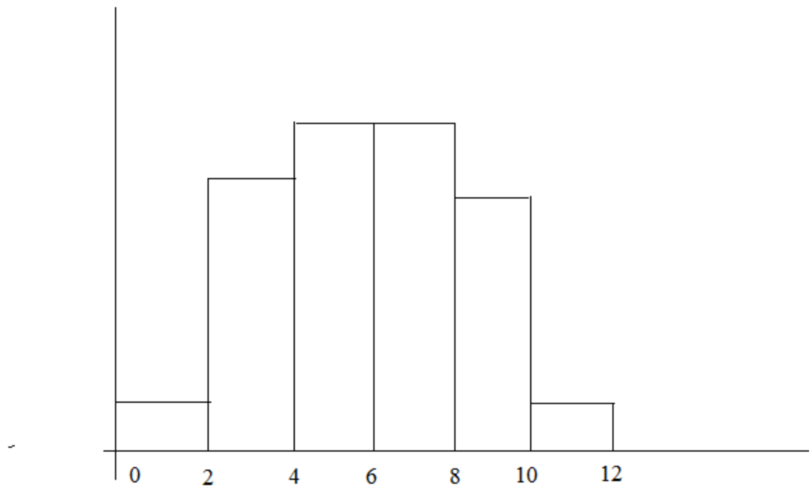


Рис. 48

Вигляд гістограми також демонструє правдоподібність гіпотези про нормальний розподіл.

3.21. Елементи теорії кореляції
Функціональна, статистична та кореляційна залежності

На практиці ми зустрічаємось з різними типами зв'язку між випадковими величинами. Одним з таких типів є *функціональна* залежність. Така залежність виникає тоді, коли зв'язок між величинами настільки тісний, що, знаючи можливі значення однієї з величин, можна точна вказати можливі значення іншої. Функціональну залежність величини Y від величини X можна задати формулою: $Y = f(X)$. Нехай, наприклад, $Y = X^2$, можливі значення величини X: 1, 2, 3. Тоді можливі значення величини Y: 1, 4, 9.

До величин, що пов'язані функціональними залежностями, відносяться значна кількість тих, що зустрічаються у природознавстві. Наприклад, тиск P, температура T та об'єм газу V пов'язані функціональною залежністю $P = RT/V$, де R – універсальна газова стала.

Разом з цим точна функціональна залежність між величинами реалізується досить рідко, оскільки величини підлягають впливу багатьох випадкових факторів. Серед цих факторів можуть бути спільні для обох величин. В таких випадках виникають статистичні залежності. Наприклад, якщо величина Y залежить від випадкових факторів Z_1, Z_2, Z_3, Z_4, а величина X – від випадкових факторів Z_3, Z_4, Z_5, то між Y та X є статистична залежність – серед випадкових факторів, що діють на величини Y та X, є спільні – Z_3, Z_4.

Означення. *Статистичною залежністю* між величинами називають таку залежність, коли зміна однієї з величин викликає зміну розподілу іншої.

Зокрема, якщо при зміні однієї з величин змінюється середнє значення іншої, то така залежність називається *кореляційною.*

Статистичні та кореляційні залежності знаходять численні застосування в усіх сферах природознавства, зокрема і в геології. Наприклад, існує кореляційна залежність між органічними речовинами, що є у ґрунті, з неорганічними мінеральними компонентами порід. Такого типу залежності лежать в основі багатьох методів підрахунку природничих ресурсів, на них спираються прогностичні побудови, розрахункові схеми.

Кореляційні зв'язки є частинним випадком статистичних. Термін «кореляція» походить від англійського *correlation* – співвідношення, відповідність. При таких залежностях певному значенню однієї з величин може відповідати одразу декілька значень іншої.

Для дослідження кореляційного зв'язку між величинами часто використовується математична модель, яка називається *рівнянням регресії*. При цьому дослідження складається з двох етапів:

1) виявлення на підставі великої кількості спостережень того, як змінюється в середньому величина Y в залежності від величини X, тобто знаходження рівняння зв'язку між Y та X;

2) виявлення ступені взаємозв'язку явищ, що досліджуються.

3.22. Побудова прямої лінії регресії методом найменших квадратів

Припустимо, що експериментально встановлено, що залежність величини Y від величини X близька до лінійної:

$$Y = \alpha X + \beta, \tag{*}$$

але параметри α і β ми не знаємо.

Нехай $x_1, x_2, ..., x_n$ – можливі значення величини X, а $y_1, y_2, ..., y_n$ – можливі значення величини Y (для спрощення припускаємо, що кожне значення зустрічається по одному разу).

Якщо рівняння (*) виконується хоча б наближено, то значення $y_1, y_2, ..., y_n$ не повинні сильно відрізнятися відповідно від значень $\alpha x_1 + \beta, \alpha x_2 + \beta, \ldots,$ $\alpha x_n + \beta$. Тому є сенс ввести міру відмінності цих значень та обрати параметри α і β таким чином, щоб ця міра відмінності була по можливості найменшою. Такою мірою, зокрема, може бути функція двох змінних:

$$S(\alpha, \beta) = \sum_{i=1}^{n} (\alpha x_i + \beta - y_i)^2 .$$

Підберемо α і β з умови мінімуму функції $S(\alpha, \beta)$. Для цього дорівняємо до нуля частинні похідні цієї функції:

$$\frac{\partial S}{\partial \alpha} = 2\sum_{i=1}^{n} (\alpha x_i + \beta - y_i) x_i = 0,$$

$$\frac{\partial S}{\partial \beta} = 2\sum_{i=1}^{n} (\alpha x_i + \beta - y_i) = 0 .$$

Перепишемо ці рівняння так:

$$\left.\begin{array}{c} \left(\sum_{i=1}^{n} x_i^2\right)\alpha + \left(\sum_{i=1}^{n} x_i\right)\beta = \sum_{i=1}^{n} x_i y_i, \\ \left(\sum_{i=1}^{n} x_i\right)\alpha + n\beta = \sum_{i=1}^{n} y_i. \end{array}\right\} \tag{**}$$

Тобто отримали систему двох лінійних алгебраїчних рівнянь відносно α і β. Систему (**) називають *системою методу найменших квадратів*. Розв'язуючи цю систему, знаходимо:

$$\alpha = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2},$$

$$\beta = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2}.$$

3.23. Вибірковий коефіцієнт кореляції

Перетворимо формули для α і β наступним чином:

$$\alpha = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}_B\overline{y}_B}{n\frac{1}{n}\left(\sum_{i=1}^{n} x_i^2\right) - n\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2} =$$

$$= \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}_B\overline{y}_B}{n\sigma_x^2} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}_B\overline{y}_B}{n\sigma_x\sigma_y}\cdot\frac{\sigma_y}{\sigma_x}.$$

Величина

$$r_B = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}_B\overline{y}_B}{n\sigma_x\sigma_y}$$

називається *вибірковим коефіцієнтом кореляції*. З використанням його отримаємо:

$$\alpha = r_B \cdot \frac{\sigma_y}{\sigma_x}.$$

Аналогічно дістанемо:

$$\beta = \overline{y}_B - r_B\frac{\sigma_y}{\sigma_x}\overline{x}_B.$$

Підставимо ці значення у рівняння прямої лінії регресії $y = \alpha x + \beta$:

$$y = r_B \frac{\sigma_y}{\sigma_x} x + \overline{y}_B - r_B \frac{\sigma_y}{\sigma_x} \overline{x}_B .$$

Або:

$$y - \overline{y}_B = r_B \frac{\sigma_y}{\sigma_x} (x - \overline{x}_B) . \qquad (*)$$

Рівняння (*) називається *вибірковим рівнянням прямої лінії регресії* Y *на* X.

Аналогічно знаходиться вибіркове рівняння прямої лінії регресії X на Y:

$$x - \overline{x}_B = r_B \frac{\sigma_x}{\sigma_y} (y - \overline{y}_B). \qquad (**)$$

На площині Oxy обидві прямі (*) та (**) проходять через одну й ту ж точку $(\overline{x}_B, \overline{y}_B)$, яка називається *центром сумісного розподілу*.

Вибірковий коефіцієнт кореляції r_B характеризує тісноту зв'язку між величинами X та Y. Нескладно довести, що $|r_B| \le 1$. Чим ближче $|r_B|$ до 1, тим тісніше зв'язок між X та Y. У цьому випадку рівняння (*) буде близьким до рівняння (**) (у випадку $r_B = \pm 1$ ці рівняння співпадають), тобто прямі лінії регресії, що описуються цими рівняннями, будуть відрізнятися незначно, кут φ між ними буде невеликим. Дійсно, тангенс цього кута виражається величиною:

$$\operatorname{tg} \varphi = \frac{\sigma_y}{\sigma_x} \cdot \frac{1 - r_B^2}{r_B},$$

звідки видно, що у випадку $|r_B| \approx 1$ буде $\operatorname{tg} \varphi \approx 0$, тобто $\varphi \approx 0$ або $\varphi \approx \pi$.

Для задач геології зв'язок між величинами вважається суттєвим, якщо $|r_B| \ge 0{,}65$.

3.24. Приклад побудови прямих ліній регресії

Нехай величини X та Y приймають значення:

X	-7,00	-6,00	-5,00	-4,00	-3,00	-2,00	-1,00	0,00	1,00
Y	7,75	6,27	5,67	4,26	3,09	2,03	1,32	0,79	-0,70

Треба знайти вибірковий коефіцієнт кореляції, визначити тісноту зв'язку між величинами X та Y, побудувати рівняння прямих ліній регресії Y на X та X на Y, побудувати ці прямі на координатній площині.

Для знаходження вибіркового коефіцієнта кореляції скористаємось формулою:

$$r_B = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}_B\overline{y}_B}{n\sigma_x\sigma_y}.$$

Знайдемо:

$$\sum_{i=1}^{n} x_i = -27, \quad \sum_{i=1}^{n} y_i = 30{,}48, \quad \sum_{i=1}^{n} x_i^2 = 141, \quad \sum_{i=1}^{n} y_i^2 = 166{,}1974, \quad \sum_{i=1}^{n} x_i y_i = -152{,}61,$$

$$\overline{x}_B = \frac{1}{n}\sum_{i=1}^{n} x_i = -3, \quad \overline{y}_B = \frac{1}{n}\sum_{i=1}^{n} y_i \approx 3{,}39,$$

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}_B^2} \approx 2{,}58, \quad \sigma_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \overline{y}_B^2} \approx 2{,}64.$$

Отже:

$$r_B \approx -0{,}996, \; |r_B| \approx 0{,}996.$$

Ми бачимо, що $|r_B|$ досить близький до 1, отже зв’язок між величинами X та Y значний.

Рівняння прямої лінії регресії Y на X:

$$y - \overline{y}_B = r_B\frac{\sigma_y}{\sigma_x}(x - \overline{x}_B).$$

Підставляючи знайдені величини, отримаємо:

$$y - 3{,}39 = -0{,}996 \cdot \frac{2{,}64}{2{,}58}(x+3).$$

Або:

$$y - 3{,}39 = -1{,}019(x+3).$$

Або:

$$y = -1{,}019x + 0{,}333.$$

Рівняння прямої лінії регресії X на Y:

$$x - \overline{x}_B = r_B\frac{\sigma_x}{\sigma_y}(y - \overline{y}_B).$$

Підставляючи знайдені величини, отримаємо:

$$x + 3 = -0{,}996 \cdot \frac{2{,}58}{2{,}64}(y - 3{,}39).$$

Або:

$$x + 3 = -0{,}973(y - 3{,}39).$$

Або:

$$y = -1{,}02x + 0{,}31.$$

Ми бачимо, що кутові коефіцієнти цих прямих майже співпадають, і обидві ці прямі проходять через точку (-3; 3,39) – центр сумісного розподілу. Отже, ці прямі також майже співпадають, що також свідчить про тісний зв'язок між величинами X та Y.

Прямі лінії регресії Y на X та X на Y побудовано на рис. 49.

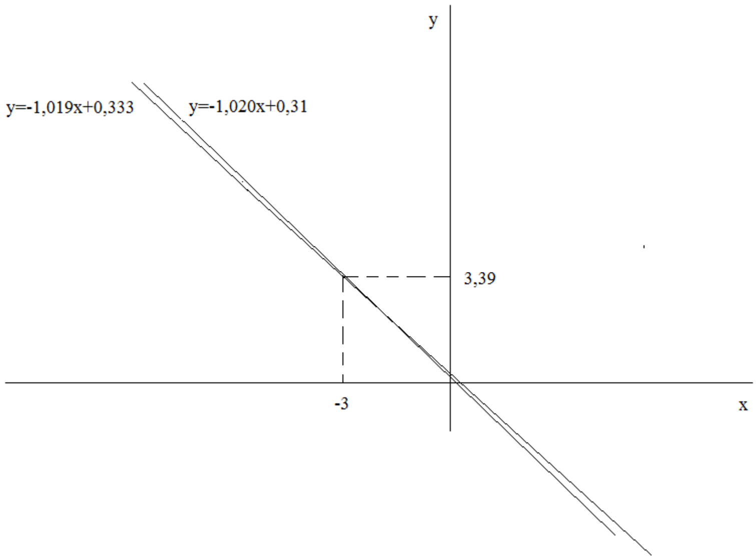


Рис. 49

Контрольні питання до розділу III

1. У чому полягає предмет математичної статистики, які її основні задачі?
2. Що є основним методом математичної статистики?
3. Що таке генеральна та вибіркова сукупності?
4. Що таке повторна та безповторна вибірка?
5. Які існують способи відбору з генеральної сукупності?
6. Що таке статистичний розподіл вибірки? Які у нього спільні риси із законом розподілу дискретної випадкової величини?
7. Що таке емпірична функція розподілу? Які у неї спільні риси з функцією розподілу випадкової величини?
8. Що таке полігон і гістограма? Як будується гістограма? Що є аналогом гістограми в теорії випадкових величин?
9. Що таке точкові оцінки параметрів розподілів, які існують вимоги до них?
10. Що таке вибіркове середнє та вибіркова дисперсія? Аналогом яких характеристик випадкових величин є ці числові характеристики вибірки?
11. Що таке початкові та центральні емпіричні моменти?
12. У чому полягає метод моментів точкових оцінок параметрів розподілів?
13. У чому полягає метод найбільшої правдоподібності точкових оцінок параметрів розподілів?
14. Що таке мода, медіана, розмах варіювання?
15. У чому полягає ідея інтервальних оцінок параметрів розподілів? У яких випадках доцільніше використання точкових оцінок, а у яких – інтервальних?
16. Що таке довірчий інтервал і довірча ймовірність?
17. Як будується довірчий інтервал для оцінки математичного сподівання нормального розподілу?
18. Що таке статистична гіпотеза?
19. Що таке критерій узгодженості? Які бувають критерії узгодженості? Як ними користуються на практиці?
20. Що таке функціональна, статистична та кореляційна залежності? Наведіть відповідні приклади.
21. Що таке вибірковий коефіцієнт кореляції, що він характеризує?
22. У чому полягає ідея методу найменших квадратів побудови прямих ліній регресії?

Вправи для самостійного розв'язування

1. Знайти емпіричну функцію заданого статистичного розподілу та побудувати її графік.

x_i	15	20	25	30
n_i	10	15	30	20

2. Побудувати полігон частот та відносних частот заданого розподілу вибірки

x_i	4	7	8	12	15	20
n_i	12	2	8	8	7	3

3. Побудувати гістограму частот за заданим розподілом вибірки:

Частинні проміжки $(h=4)$	Сума частот вибірки n_i^*	Щільність частот n_i^*/h
1 – 5	10	2,5
5 – 9	20	5
9 – 13	50	12,5
13 – 17	12	3
17 –21	8	2

4. З генеральної сукупності добуто вибірку:

x_i	1	3	6	26
n_i	8	40	10	2

Знайти незсунуту оцінку середнього генерального.

5. Знайти вибіркову дисперсію та вибіркове середньоквадратичне відхилення за даним розподілом вибірки:

x_i	2502	2804	2903	3028
n_i	8	30	6	2

6. Знайти вибіркову дисперсію та вибіркове середньоквадратичне відхилення за даним розподілом вибірки:

x_i	18,4	18,9	19,3	19,6
n_i	5	10	20	15

7. Знайти виправлену вибіркову дисперсію за даним розподілом вибірки:

x_i	0,01	0,05	0,09
n_i	2	3	5

8. Вибірка з великої партії електроламп містить 100 ламп. Середня тривалість горіння однієї лампи 1000 годин. Знайти з надійністю $\gamma = 0,95$ довірчий інтервал для середньої тривалості a горіння всієї партії, якщо відомо, що середньоквадратичне відхилення тривалості горіння лампи $\sigma = 40$ годин.

9. З генеральної сукупності добуто вибірку:

x_i	-2	1	2	3	4	5
n_i	2	1	2	2	2	1

З надійністю $\gamma = 0,95$ знайти довірчий інтервал для оцінки математичного сподівання нормально розподіленої ознаки генеральної сукупності.

10. З генеральної сукупності добуто вибірку:

x_i	-0,5	-0,4	-0,2	0	0,2	0,6	0,8	1	1,2	1,5
n_i	1	2	1	1	1	1	1	1	2	1

З надійністю $\gamma = 0{,}95$ знайти довірчий інтервал для оцінки середньоквадратичного відхилення нормально розподіленої ознаки генеральної сукупності.

11. За допомогою критерія Пірсона на рівні значущості 0,05 перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо з неї добуто вибірку:

x_i	5	7	9	11	13	15	17	19	21
n_i	15	26	25	30	26	21	24	20	13

Додаток

Таблиця 1

Значення функції $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3726	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001
4,0	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001

Таблиця 2

Значення функції $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$

	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0.4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0.4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998

Таблиця 3

Значення $t_\gamma = t_\gamma(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Таблиця 4

Значення $q = q(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1.62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	3,97	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Таблиця 5

Критичні точки χ^2_{kp} критерія Пірсона

Число ступенів вільності k	Рівень значущості ά 0,01	0,025	0,05	0,1
1	6,635	5,024	3,841	2,706
2	9,210	7,378	5,991	4,605
3	11,345	9,348	7,815	6,251
4	13,277	11,143	9,488	7,779
5	15,086	12,832	11,070	9,236
6	16,812	14,449	12,592	10,645
7	18,475	16,013	14,067	12,017
8	20,090	17,535	15,507	13,362
9	21,666	19,023	16,919	14,684
10	23,209	20,483	18,307	15,987
11	24,725	21,920	19,675	17,275
12	26,217	23,337	21,026	18,549
13	27,688	24,736	22,362	19,812
14	29,141	26,119	23,685	21,064
15	30,578	27,488	24,996	22,307
16	32,000	28,845	26,296	23,542
17	33,409	30,191	27,587	24,769
18	34,805	31,526	28,869	25,989
19	36,191	32,852	30,144	27,204
20	37,566	34,170	31,410	28,412
21	38,932	35,479	32,676	29,615
22	40,289	36,781	33,924	30,813
23	41,638	38,076	35,172	32,007
24	42,980	39,364	36,415	33,196
25	44,314	40,646	37,652	34,382
26	45,642	41,923	38,885	35,563
27	46,963	43,194	40,113	36,741
28	48,278	44,461	41,337	37,916
29	49,588	45,722	42,557	39,087
30	50,892	46,979	43,773	40.256

Список рекомендованої літератури

Базова

1. Гмурман В.Е. Теория вероятностей и математическая статистика. М.,1977.
2. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. М., 1988.
3. Сеньо П. С. Теорія ймовірностей та математична статистика. К., 2004.
4. Барковський В.В., Барковська Н.В., Лопатін О.К. Теорія ймовірностей та математична статистика. К., 2002.
5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М., 1975.
6. Агапов Г. Е. Задачник по теории вероятностей. М., 1994.

Додаткова

1. Гнеденко Б. В. Курс теории вероятностей. М., 1954.
2. Гихман И. И., Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика. К., 1979.
3. В. Феллер. Введение в теорию вероятностей и её приложения. Т1, Т.2. М.,1967.
4. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. М., 1983.

Зміст

Навчальне видання

Щоголев Сергій Авенірович

ОСНОВИ ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

НАВЧАЛЬНО-МЕТОДИЧНИЙ ПОСІБНИК

За редакцією автора

Підп. до друку 03.09.2015. Формат 60х84/16.
Умов.-друк. арк. 11,86. Тираж 40 пр.
Зам. № 1209.

Видавець і виготовлювач
Одеський національний університет
імені І. І. Мечникова
Свідоцтво суб’єкта видавничої справи ДК № 4215 від 22.11.2011 р.

Україна, 65082, м. Одеса, вул. Єлісаветинська, 12
Тел.: (048) 723 28 39. E-mail: druk@onu.edu.ua