В.Т. Швець

Теорія ймовірностей, математична статистика та випадкові процеси

Одеса 2021

ББК 22.161я73
УДК 517(0.75.8)

Швець Валерій Тимофійович
Теорія ймовірностей, математична статистика та випадкові процеси
Одеса.Видавництво (В електронному вигляді),2021 - 234 с.

Навчальний посібник є вступом у теорію ймовірностей, математичну статистику та випадкові процеси і, в цілому, відповідає програмі відповідної дисципліни технічних вузів. Від підручників для класичних університетів він відрізняється ширшим охопленням проблематики теорії ймовірносте, математичної статистики та випадкових процесів і меншою строгістю викладення матеріалу. Доводяться лише засадничі для теорії теореми. Підбір матеріалу для навчального посібника зорієнтований на подальше використання для обробки експериментальних даних різноманітних експериментів. Посібник містить велику кількість практичних задач і їх розв'язань, так само як і велику кількість задач, призначених для самостійної роботи. Це робить його зручним для використання студентами у всіх формах навчальної роботи з вищої математики. Посібник розрахований на студентів та аспірантів технічних вузів.

Автор:
доктор фіз.-мат. наук, проф. В.Т. Швець
Рецензенти:
доктор фіз.-мат. наук, проф. С. В. Козицький
доктор фіз.-мат. наук, проф. Я. І. Лепіх
доктор тех. наук, проф. О. В. Дорошенко

Зміст

Передмова

Традиційно в математичній підготовці студентів технічних вузів основну увагу приділяють таким дисциплінам як лінійна алгебра, аналітична геометрія та математичний аналіз. Ці дисципліни, в першу чергу, потрібні в якості підвалин всієї математичної науки і досить віддалені від реального життя. Наше життя просякнуте не стільки строго детермінованими закономірностями, скільки такими, що мають ймовірнісний характер. "Все, що відбулось, не могло не відбутись. Все, що не відбулось, не могло відбутись". таким є засадничий принцип всіх світових релігій, що відповідає становленню засадничих уявлень людини про навколишній світ на початку розвитку людської цивілізації. Майбутнє, як правило, існує у багатьох варіантах. Підкиньте монету і Ви в цьому переконаєтесь. Сучасний світ людини наповнений процесами, що описуються статистичними закономірностями. Квантова механіка, що описує поведінку поля і частинок на мікроскопічному рівні принципово ґрунтується на описі цих об'єктів мовою ймовірностей, функцій розподілу, математичного очікування тощо. Все це свідчить про те, що роль таких наук в математичній підготовці студентів технічних вузів як: теорія ймовірностей, математична статистика, випадкові процеси тощо має невпинно зростати. Аудиторія тих, кому потрібні ці знання, також невпинно зростає. Проте велика кількість підручників, що існують з названих вище дисциплін, призначені для студентів університетів, а можливо навіть для викладачів математичних факультетів університетів. Зробити знання з теорії ймовірностей доступними не лише студентам мехмату але і широкому загалу студентів технічних вузів - ось велике завдання, що стоїть перед викладачами вузів. Ця задача полегшується тим, що засадничі принципи теорії ймовірностей і інших, пов'язаних з нею наук, насправді, надзвичайно прості. Переконати в цьому студентів технічних вузів і є завданням даного навчального посібника.

Глава 1. Елементи теорії ймовірностей

1.1.Частотне означення ймовірності

Найважливішим поняттям теорії ймовірностей є ймовірність. Теорія вивчає властивості цього математичного об'єкту та різні алгоритми його обчислення. Ймовірність є зручною числовою (дійсні числа) характеристикою наслідків випадкового експерименту. Вона дозволяє порівнювати ці наслідки між собою, використовуючи знаки: <, >, =. Випадковим є експеримент з неоднозначними наслідками. Наприклад, підкидання монети, постріл по мішені. В принципі, будь-які експерименти можна розглядати в рамках такого підходу включно з експериментами, що мають однозначні наслідки, як частинні випадки експериментів з неоднозначними наслідками. Наслідки випадкових експериментів прийнято називати подіями, а самі випадкові експерименти - випробуваннями. Подія, не звідна до сукупності простіших подій називається елементарною подією. Всі елементарні події попарно несумісні. Тобто, якщо відбулась одна з них, то це унеможливлює одночасне настання будь-якої іншої елементарної події. Якщо ми підкидаємо монету один раз, то можуть спостерігатись лише дві події: випадення герба або решки. Кожна з цих подій є елементарною. При підкиданні гральної кості подій вже досить багато. Шість з них можна розглядати як елементарні - це випадення чисел від 1 до 6. Випадення парного числа - це вже подія, що складається з трьох елементарних подій: випадення чисел 2, 4 або 6. У теорії ймовірностей кожній події ставиться у відповідність її числова характеристика - ймовірність. Ймовірність є абстрактним поняттям, суть якого можна зрозуміти, пов'язавши його з простішою числовою характеристикою подій, що називається частотою подій. Якщо один і той самий випадковий експеримент ми проводимо в однакових умовах декілька разів, то частота означає долю від загальної кількості випадкових експериментів (випробувань), в якій спостерігається подія, що нас цікавить. Частота легко вимірюється експериментально. За означенням частоти $\nu(A)$ події A

$$\nu_N(A) = \frac{n_N(A)}{N}.$$

Тут N - загальна кількість випробувань, $n_N(A)$ - кількість випробувань, в яких спостерігається подія A, яка, зрозуміло, також залежить від N. Означена таким чином частота, очевидно, є випадковою величиною, а не числом, оскільки залежить від N. Також вона може бути іншою для іншої серії N випробувань. В цьому її незручність, як характеристики випадкового експерименту. Разом з тим, легко експериментально переконатись у тому, що діапазон коливань значень частоти зменшується із зростанням N, якщо всі випробування відбуваються в одних і тих самих умовах. Якщо при зростанні кількості випробувань частота події A прямує до певної межі, то ця межа і називається ймовірністю події A

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

Це експериментальне означення ймовірності належить німецькому математику Ріхарду Едлеру фон Мізесу (19.04.1883 – 14.07.1953). Перевага ймовірності порівняно з частотою для характеристики наслідків випадкового експерименту у тому, що ймовірність є певним числом, а не випадковою величиною. При досить великих значеннях N частота мало відрізняється від ймовірності і їх можна практично не розрізняти. Інша математична дисципліна - математична статистика дозволяє теоретично дослідити деталі процесу наближення частоти до ймовірності. Всі властивості ймовірності є наслідком властивостей частоти. Наприклад, очевидно, що частота, а відповідно і ймовірність, змінюється в інтервалі $[0,1]$.

Далі, якщо події A і B є несумісними - не можуть відбуватись одночасно - виключають одна одну, то

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

Тут під подією $A+B$ розуміється подія, що відбувається або, коли відбувається подія A, або, коли відбувається подія B. Ця властивість - адитивність частот несумісних подій має місце для довільної кількості випробувань N. Очевидно, що властивість адитивності має місце і для ймовірності

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

оскільки

$$\lim_{N\to\infty} \nu_N(A+B) = \lim_{N\to\infty} \nu_N(A) + \lim_{N\to\infty} \nu_N(B)$$

Нарешті, якщо події A і B незалежні - між ними немає причинно-наслідкового зв'язку, то для частоти події $A*B$, що відповідає одночасному настанню подій A і B, має місце наступне наближене експериментальне співвідношення

$$\nu_N(AB) = \nu_N(A)\nu_N(B).$$

Ця властивість - мультипликативність частот незалежних подій стає тим точнішою, чим більшою є кількість випробувань. Певним додатковим припущенням, інтуїтивно прийнятним, є те, що властивість мультипликативності з наближеного для частоти перетворюється у точну рівність для ймовірності

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

також частково основану на властивості межового переходу

$$\lim_{N\to\infty} \nu_N(A)\nu_N(B) = \lim_{N\to\infty} \nu_N(A) \lim_{N\to\infty} \nu_N(B).$$

Частотне означення ймовірності є універсальним, зручним для експериментального обчислення ймовірності але непридатним для її теоретичного обчислення.

1.2.Аксіоматичне означення ймовірності

На завершальному етапі свого розвитку будь-яка математична теорія прагне набрати рис аксіоматичної теорії, першим довершеним зразком якої стала геометрія Евкліда. Тобто в основу теорії кладеться система аксіом, а вся будова теорії є наслідками цих аксіом. Цим досягається більша логічна ясність самої теорії, а також, дуже часто, більша простота її обчислювальних алгоритмів. В якості аксіом звичайно беруться твердження інтуїтивно зрозумілі, або, як це часто буває для фізичних теорій, такі, що є доконаними експериментальними фактами. У разі теорії ймовірностей, її основний об'єкт - ймовірність. Ймовірність можна ввести аксіоматично не пов'язуючи її з частотою, а постулюючи її властивості, що, насправді, випливають з властивостей частоти. Формалізацію теорії ймовірностей можна підсилити, якщо залучити до неї в якості складової частини теорію множин з її розвиненим математичним апаратом. Цього можна досягти ототожнивши подію з множиною, а елементарну подію з елементом множини. У висліді події перетворюються у добре визначені математичні об'єкти - множини, а дії над множинами стають діями над подіями.

Таким чином, ймовірність можна задати як числову функцію події (множини), що має наступні властивості:

1.Її множина значень $0 \le P(A) \le 1$. При цьому ймовірність неможливої події - пустої множини $P(\varnothing) = 0$, ймовірність достовірної події - події, підмножинами якої є всі можливі випадкові події даного випадкового експерименту, $P(\Omega) = 1$.

2.Її областю визначення є всі можливі події, що характеризують даний випадковий експеримент разом з неможливою і достовірною подіями. Множина всіх підмножин скінченої або зліченої множини завжди утворює алгебру, тобто множину, замкнену відносно всіх операцій над подіями: додавання, віднімання, множення і заперечення. Тому можна стверджувати, що областю визначення ймовірності є алгебра подій U. Якщо з якихось причин неможливо розглянути максимальну кількість подій, то фактично розглянута сукупність подій має утворювати алгебру. Найпростішою алгеброю є сукупність двох подій: неможливої і достовірної.

Розглядають також σ - алгебру подій. Це, коли множина подій додатково замкнена відносно нескінченої, але зліченої кількості операцій додавання і множення подій. Випадок незлічених множин вимагає певних виключень з твердження множина всіх підмножин, але про це пізніше.

Зауваження. Задаючи область визначення класичної функції в якості відрізку числової осі, ми задаємо значення функції лише у кожній точці цього відрізку, тобто лише для кожного окремого елементу відповідної числової множини. Задаючи область визначення такої специфічної функції, як ймовірність, ми задаємо її значення не лише для кожного елемента простору елементарних подій Ω, але і для довільних підмножин цієї множини, що утворюють алгебру.

3.Властивість адитивності ймовірності скінченої суми несумісних подій

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

Цю властивість можна підсилити, розповсюдивши її на нескінчену суму несумісних подій.

4.Властивість мультипликативності ймовірності скінченого добутку незалежних подій

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

Цю властивість також можна підсилити, поширивши її на нескінчений добуток незалежних подій. Звичайно цю властивість не включають до переліку аксіом, що визначають ймовірність, проте на певному етапі розгляду ймовірностей ця аксіома все одно з'являється.

5.Властивість неперервності ймовірності - аналог неперервності числової функції числового аргументу

$$\lim_{n\to\infty} P(A_n) = 0,$$

де A_n - спадна послідовність подій така, що кожна наступна подія є лише частиною попередньої $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ і $\prod_{n=1}^{\infty} A_n = \varnothing$.

Нульове значення ймовірності тут не є принциповим. Ця умова схожа на умову неперервності класичних функцій

$$\lim_{x\to x_0} f(x) = f(x_0),$$

яка означає, що послідовність значень функції збігається до значення функції у тій точці, яка є межовою для послідовності значень аргументу.

Аксіоматичне означення ймовірності, так само як і частотне, є універсальним, а тому не дає рецептів її конкретного обчислення. Трійку математичних об'єктів, що складається з простору елементарних подій, алгебри подій і визначеній на ній ймовірності, (Ω, U, P) прийнято називати ймовірнісним простором. Перші два елемента у дужках задають область визначення функції, якою є третій елемент – ймовірність.

Якщо простір елементарних подій складається із скінченої або зліченої кількості елементарних подій, то такий ймовірнісний простір називається дискретним. Якщо елементарні події утворюють континуум, то неперервним. Аксіома $P(\Omega) = 1$

фактично є умовою нормування ймовірності. Для дискретного ймовірнісного простору $\Omega=\{\omega_1,\omega_2,...,\omega_n,...\}$ остання має вигляд

$$P(\Omega)=P(\{\omega_1,\omega_2,...,\omega_n,...\})=$$

$$=P(\omega_1+\omega_2+...+\omega_n+...)=$$

$$=P(\omega_1)+P(\omega_2)+...+P(\omega_n)+...=1.$$

Теорема додавання ймовірностей. Нехай A і B - дві довільні події, можливо сумісні. Тоді

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

Доведення. *Теорія множин разом з аксіоматичним означенням ймовірності дозволяють виконати вишукане доведення цієї теореми. Дійсно, представимо суму двох довільних подій як суму двох несумісних подій*

$$A+B=A+B\overline{A}.$$

Тут $\overline{A}$ є запереченням події A, тобто подія не A. З іншого боку, подію B також можна представити сумою двох несумісних подій

$$B=AB+B\overline{A}.$$

Якщо перейти до ймовірностей, то останні дві рівності перетворяться за допомогою третьої аксіоми ймовірностей на наступні

$$P(A+B)=P(A)+P(B\overline{A}),$$

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

Виключаючи звідси ймовірність $P(B\overline{A})$ ми і отримаємо шуканий вираз.

Для трьох довільних подій теорема додавання ймовірностей матиме вигляд

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

$$-P(AB)-P(AC)-P(BC)+P(ABC),$$

що доводиться аналогічно попередньому. Теорему додавання ймовірностей можна узагальнювати і далі.

Теорема множення ймовірностей. Нехай A і B - дві довільні події, можливо залежні. Тоді

$$P(AB) = P(A)P(B \mid A).$$

Ця формула не доводиться. Вона є узагальненням четвертої аксіоми ймовірності у разі залежних подій і, фактично, є означенням функції $P(B \mid A)$ - умовної ймовірності події B, тобто ймовірності події B при умові, що подія A відбулась. У разі трьох довільних подій ця формула матиме вигляд

$$P(ABC) = P(A)P(B \mid A)P(C \mid AB).$$

Тут $P(C \mid AB)$ є ймовірністю події C при умові, що відбулись події A і B. Подальше узагальнення теореми множення ймовірностей не викликає ускладнень.

1.3.Елементи комбінаторики

Є цілий ряд простих і надзвичайно важливих з практичної точки зору випадків, коли аксіоматичне означення ймовірності приводить до простих алгоритмів її обчислення. Найпростіші алгоритми зводяться до простого підрахунку кількості елементарних подій, з яких складається дана подія, і кількості елементарних подій, з яких складається простір елементарних подій. Часто доводиться мати справу з підрахунком великої кількості елементарних подій. Розділ математики під назвою комбінаторика полегшує цей підрахунок, тому розглянемо її основи.

Комбінаторикою називається розділ математики, присвячений вибору та розміщенню елементів множини.

Упорядкованою множиною називають довільну множину, де враховано порядок розміщення її елементів. Довільну n – елементну впорядковану множину можна позначити так

$$A = \left(x_1, x_2, \ldots, x_n \right).$$

Якщо порядок елементів ролі не грає, то множина називається невпорядкованою. Довільну n – елементну невпорядковану множину можна позначити так

$$B = \left\{ y_1, y_2, \ldots, y_n \right\}.$$

Упорядковані множини називаються рівними, якщо вони складаються з однакових елементів, розташованих в однаковому порядку. Так множини $A = \left(1,2\right)$ і $B = \left(1,2\right)$ є впорядкованими рівними множинами, а $A = \left(1,2\right)$ і $C = \left(2,1\right)$ є впорядкованими різними множинами, тобто

$$A = B, \; A \neq C.$$

Невпорядковані множини називаються рівними, якщо складаються з однакових елементів. Так множини $A = \left\{1,2,3\right\}$ і $B = \left\{2,1,3\right\}$ є невпорядкованими рівними множинами, а $A = \left\{1,2,3\right\}$ і $C = \left\{1,3,4\right\}$ є невпорядкованими різними множинами, тобто

$$A = B, \; A \neq C.$$

Розміщенням з n елементів по m називають упорядковану m – елементну множину, утворену з невпорядкованої n – елементної множини. Кількість усіх таких розміщень A_n^m визначається формулою

$$A_n^m = n\left(n-1\right)\ldots\left(n-m+1\right) = \frac{n!}{\left(n-m\right)!}, \; 0 \leq m \leq n,$$

де $0! = 1$, $n! = 1 \cdot 2 \ldots n$, $n \in N$. Зокрема, $A_n^0 = 1$, $A_0^0 = 1$.

Якщо $M = \{1,2,3\}$ – триелементна невпорядкована множина, то з її елементів можна утворити $A_3^1 = 3$ розміщень з трьох по одному елементу: (1), (2), (3); $A_3^2 = 3 \cdot 2 = 6$ розміщень з трьох по два елементи: $(1,2)$, $(2,1)$, $(1,3)$, $(3,1)$, $(2,3)$, $(3,2)$; $A_3^3 = 3 \cdot 2 \cdot 1 = 6$ розміщень з трьох по три елементи: $(1,2,3)$, $(3,1,2)$, $(2,3,1)$, $(2,1,3)$, $(3,2,1)$, $(1,3,2)$.

Перестановкою n – елементної множини називають розміщення з n – елементів по n елементам. Кількість таких перестановок обчислюється за формулою

$$P_n = A_n^n = n!.$$

Сполученням з n – елементів по m називають довільну невпорядковану m елементну множину, утворену з невпорядкованої n – елементної множини. Кількість сполучень визначається формулою

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n!}{m!(n-m)!} = \frac{n(n-1)\ldots(n-m+1)}{m!},\ 0 \le m \le n.$$

Зокрема, $C_n^0 = 1$, $C_n^n = 1$.

Якщо $M = \{1,2,3\}$ – триелементна невпорядкована множина, то з її елементів можна створити $C_3^1 = 3$ сполучень з трьох елементів по одному: $\{1\}$, $\{2\}$, $\{3\}$; $C_3^2 = 3$ сполучень з трьох елементів по два: $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $C_3^3 = 1$ сполучень з трьох елементів по три: $\{1,2,3\}$.

Вибірка. Довільний набір з елементів множини M, які можуть повторюватись, називається вибіркою, або рядом. Кількість елементів вибірки називається її довжиною. Вибірку довжиною k можна позначити так $(x_1, x_2, \ldots x_k)$, де $x_i \in M$, $i \in N$

. Довжина вибірки може перевищувати кількість елементів множини M.

Нехай $M=\{1,2,3\}$, тоді з елементів цієї множини можна утворити безліч вибірок, наприклад, (1), $(1,1)$, $(1,1,1)$, ..., (2), $(2,2)$, $(2,2,2)$, ..., (3), $(3,3)$, $(3,3,3)$, ..., $(1,2)$, $(1,1,2)$, $\ldots$.

Вибірки можуть бути впорядкованими, коли враховується порядок розміщення їх елементів, або невпорядкованими, коли цей порядок не враховується.

Дві впорядковані вибірки називаються рівними, якщо вони мають однакові довжини та однакові елементи на однакових місцях. Дві невпорядковані вибірки називаються рівними, якщо вони мають однакові довжини і кожний їх елемент повторюється однакову кількість разів. В іншому разі вибірки вважаються різними.

Розглянемо дві вибірки $(1,2,1)$ і $(1,1,2)$. Якщо їх вважати впорядкованими, то вони різні. Якщо їх вважати невпорядкованими, то вони однакові (рівні).

Розміщеннями з повтореннями з n елементів по m називають упорядковані вибірки довжиною m, утворені з n – елементної множини M. Кількість таких розміщень з повтореннями визначають за формулою

$$\bar{A}_n^m = n^m.$$

Якщо $M=\{1,2,3\}$, то розміщеннями з повтореннями з 3 – х елементів по 2 є такі множини: $(1,1)$, $(1,2)$, $(1,3)$, $(2,1)$, $(2,2)$, $(2,3)$, $(3,1)$, $(3,2)$, $(3,3)$. Їх загальна кількість $\bar{A}_3^2=3^2=9$.

Перестановками з повтореннями з n елементів по m називають упорядковані вибірки довжиною m, утворені з n – елементної множини n, при умові що $m=m_1+m_2+\ldots+m_n$, де m_1

– кількість елементів a_1 вибірки, m_2 – кількість елементів a_2 вибірки, тощо. $M = \{a_1, a_2, a_n\}$. Кількість перестановок з повтореннями визначається формулою

$$P_m = \frac{m!}{m_1!m_2!\ldots m_n!}.$$

Множина $M = \{1,2\}$ породжує такі перестановки з повтореннями з 2 – х елементів по 3: $(1,1,1)$, $(2,2,2)$, $(1,1,2)$, $(1,2,1)$, $(2,1,1)$, $(2,2,1)$, $(2,1,2)$, $(1,2,2)$. При цьому, якщо $m_1 = 3$, $m_2 = 0$, то це лише одна перестановка $P_3 = 3!/3!\cdot 0! = 1$. Якщо $m_1 = 0$, $m_2 = 3$, то це також одна перестановка $P_3 = 3!/0!3!$. Якщо ж $m_1 = 2$, $m_2 = 1$, або $m_1 = 1$, $m_2 = 2$, то це вже три перестановки $P_3 = 3!/2!\cdot 1! = 3$.

Сполученнями з повтореннями з n елементів по m називають невпорядковану вибірку довжиною m, утворену з n – елементної множини M. Кількість таких сполучень з повтореннями обчислюють так

$$C_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}.$$

Нехай $M = \{1,2,3\}$. З цієї множини можна утворити наступні сполучення з 3 – х елементів по 2: $(1,1)$, $(1,2)$, $(1,3)$, $(2,2)$, $(2,3)$, $(3,3)$. Їх загальна кількість $C_3^2 = C_4^2 = \dfrac{4!}{2!2!} = 6$.

Правило сум. Якщо деякий вибір A можна здійснити m різними способами, а вибір B – n різними способами, то вибір A або B можна здійснити $n+m$ способами.

Правило добутку. Якщо деякий вибір можна здійснити m різними способами, і після кожного з них вибір B можна здійснити n способами, то вибір A і B можна здійснити $m \cdot n$ способами.

Задача 1. *Студенти вивчають* 8 *навчальних дисциплін. Скількома способами можна скласти розклад занять на понеділок, якщо в цей день треба запланувати три лекції з різних предметів.*

Розв`язання. *Кількість таких способів дорівнює числу розміщень з* 8 *елементів по* 3, *тобто* $A_8^3 = 8 \cdot 7 \cdot 6 = 336$.

Задача 2. *У шаховому турнірі, де учасники зустрічаються між собою один раз, три шахісти вибули через хворобу, зігравши відповідно одну, дві та три партії (з учасниками, що не вибули). Скільки шахістів розпочали турнір, якщо всього були зіграні* 84 *партії?*

Розв`язання. *Позначимо через* n *число учасників турніру. Оскільки три з них вибули, зігравши в сумі* 6 *партій, то в останніх* $84 - 6 = 78$ *партіях взяли участь* $n - 3$ *учасника. Отже,* $78 = \mathrm{C}_{n-3}^2$, *тобто*

$$\frac{(n-3)(n-4)}{2} = 78 .$$

Звідси

$$n = \frac{7}{2} \pm \sqrt{\frac{49}{4} + 144} = \frac{7}{2} \pm \frac{25}{2}.$$

Взявши лише додатний корінь, матимемо $n = 16$.

Задача 3. *У круглому столі приймає участь шість осіб. Скількома способами можна розмістити учасників. Скільки буде таких варіантів, якщо дві конкретні особи мають сидіти поряд?*

Розв`язання. *У першому варіанті розміщення кількість варіантів така* $P_6 = 6! = 720$. *Розглянемо тепер другий варіант* *розміщення. Одну зазначену особу можна розмістити 6 – ма способами. Дві зазначені особи можна розмістити* $6 \cdot 2 = 12$ *способами, що враховують варіант, коли перша особа сидить ліворуч від другої, та перша особа сидить праворуч від другої. На місцях, що залишились, решту учасників можна розмістити* $P_4 = 4! = 24$. *Загальна кількість способів для всіх учасників буде* $12 \cdot 4! = 288$.

1.4.Класичне означення ймовірності

Можливість безпосереднього обчислення ймовірності обмежується декількома простими, хоча і практично важливими, випадками. Один з таких випадків описується класичним означенням ймовірності. Практично вся практика використання азартних ігор у світі базована на класичному означенні ймовірності, як на математичної моделі цього бізнесу. Це означення надає нам алгоритм обчислення ймовірності у разі, коли всі елементарні події рівноправні і нам лише доводиться підраховувати їх кількості.

Розглянемо простір елементарних подій $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$ *, що містить скінчену кількість елементарних подій. Очевидно, що з умови нормування і умови адитивності ймовірності для несумісних подій (всі елементарні події є несумісними)*

$$P(\Omega) = P(\omega_1) + P(\omega_2) + ... + P(\omega_N) = 1$$

Нехай всі елементарні події рівноправні, тобто

$$P(\omega_1) = P(\omega_2) = ... = P(\omega_n) = P(\omega).$$

Тоді умову нормування можна записати так

$$NP(\omega) = 1.$$

Звідси знаходимо елементарну ймовірність

$$P(\omega) = 1/N.$$

Таким чином, для рівноправних подій для знаходження елементарної ймовірності досить лише порахувати кількість елементарних подій в просторі елементарних подій.

Тепер розглянемо ймовірність події A. *Нехай ця подія складається з наступних елементарних подій, що є певною підмножиною простору елементарних подій*

$$A = \{\omega_{i1}, \omega_{i2}, ..., \omega_{in}\}.$$

З тих же міркувань, що і вище

$$P(A) = P(\omega_{i1}) + P(\omega_{i2}) + ... + P(\omega_{in}).$$

Оскільки всі елементарні події рівноправні, а у висліді всі елементарні ймовірності рівні, то

$$P(A) = nP(\omega) = n / N.$$

Позначаючи кількість елементів скінченої множини A як $|A|$, а кількість елементарних подій простору елементарних подій Ω як $|\Omega|$, останній результат можна записати так

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

Фактично класичне означення ймовірності відповідає випадку, коли множині можна поставити у відповідність її найпростішу міру – кількість елементів.

Розглянемо декілька прикладів.

Задача 1. *Дитина, граючись десятьма кубиками, на яких написані літери* M, M, T, T, A, A, A, K, $И$, E, *склала слово "математика". Чи можна вважати дитину освіченою?*

Розв`язання. *Елементарною подією вважатимемо довільне розташування кубиків. При цьому кубики з однаковими літерами вважатимемо різними. Кількості елементарних подій, що утворюють простір елементарних, відповідає кількість можливих перестановок десяти кубиків* $|\Omega| = P_{10} = 10!$. *Кубики з літерами, що зустрічаються в слові один раз, мають стояти на* *певних місцях. Кубик з літерою* A *можна розташувати на трьох місцях* $P_3 = 3!$ *способами; кубики з літерою* M *можна розташувати на двох місцях* $P_2 = 2!$ *способами, так само як і кубики з літерою* T – $P_2 = 2!$ *способами. Тоді подія* A*, що відповідає утворенню з літер на кубиках слова математика, складається з* $|A| = 3!\cdot 2!\cdot 2! = 24$ *подій. Відповідно до класичного означення ймовірності*

$$P\left(A\right) = \frac{|A|}{|\Omega|} = \frac{24}{10!} = \frac{1}{15120}.$$

Ця ймовірність надзвичайно мала і таку подію A *можна вважати практично неможливою. Якщо вона здійснилась, то слід вважати гіпотезу про неосвіченість дитини помилковою.*

Задача 2. *У людини в кишені* n *ключів, з яких тільки один підходить до його дверей. Ключі послідовно витягуються без повернення до кишені до тих пір, поки не з'явиться потрібний ключ. Знайти ймовірність того, що потрібний ключ з'явиться при* k *– ому витягуванні.*

Розв`язання. *Елементарною подією вважатимемо довільну послідовність з* n *ключів. Кількість таких послідовностей, тобто кількість елементарних подій, з яких складається простір елементарних подій, очевидно, є* $|\Omega| = P_n = n!$. *Послідовностей, у яких потрібний ключ знаходиться на певному місці є* $P_{n-1} = \left(n-1\right)!$*, оскільки одне місце зайняте потрібним ключем, а решта* $n-1$ *ключів можна на* $n-1$ *місцях розкласти* $|A| = P_{n-1} = \left(n-1\right)!$ *способами. Таким чином шукана ймовірність дорівнює*

$$P\left(A\right) = \frac{|A|}{|\Omega|} = \frac{P_{n-1}}{P_n} = \frac{\left(n-1\right)!}{n!} = \frac{1}{n}.$$

Задача 3. *З ящика, що містить три білети з номерами 1, 2, 3, виймають по одному всі білети. Припускаючи, що всі послідовності номерів білетів мають однакові ймовірності, знайти ймовірність того, що хоча б у одного білета порядковий номер збігається з власним.*

Розв`язання. *Нехай місця розташування білетів пронумеровані, тоді елементарними подіями вважатимемо наступні розташування білетів:*
$\omega_1 = 1,2,3$, $\omega_2 = 2,3,1$, $\omega_3 = 3,1,2$, $\omega_4 = 3,2,1$, $\omega_5 = 2,1,3$, $\omega_6 = 1,3,2$. *Простір елементарних подій* $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. *Загальна кількість елементарних подій* $|\Omega| = 6$. *Подія* $A = \{\omega_1, \omega_4, \omega_5, \omega_6\}$, *що нас цікавить, складається з чотирьох елементарних подій* $|A| = 4$. *Відповідно до класичного означення ймовірності*

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{3}.$$

Задача 4. *У колоді 36 карт. Знайти ймовірності подій:* $A =$ *{чотири тузи розташовані поряд},* $B =$ *{місця розміщення тузів утворюють арифметичну прогресію з кроком 7}.*

Розв'язання. *Елементарною подією вважатимемо довільне розташування карт у колоді. Простір елементарних подій складатиметься з* $|\Omega| = 36!$ *подій. Нехай спочатку чотири тузи розташовані поряд на початку колоди. Кількість елементарних подій, що відповідають такому розташуванню тузів дорівнюватиме добутку всіх можливих перестановок тузів між собою на кількість перестановок між собою решти карт. Зсуваючи по колоді чотири тузи від її початку що кінця, матимемо у 33 рази більше елементарних подій. Таким чином,* $|A| = 4!32!33$. *У висліді*

$$P(A)=\frac{4!32!33}{36!}=\frac{1}{3\cdot 35\cdot 17}=\frac{1}{1785}=0.00056\,.$$

Аналогічно

$$P(B)=\frac{4!32!15}{36!}=\frac{4\cdot 3\cdot 2\cdot 15}{36\cdot 35\cdot 34\cdot 33}=\frac{1}{3\cdot 17\cdot 7\cdot 11}=\frac{1}{3927}=0.00026\ .$$

Задача 5. *Кинуті три монети. Вважаючи всі елементарні події рівноправними, знайти ймовірності подій: $A=${перша монета випала гербом догори}, $B=${випало рівно два герба}, $C=${випало не більше двох гербів}.*

Розв'язання. *В якості елементарної події візьмемо довільний результат підкидання трьох монет, Очевидно загальна кількість елементарних подій $|\Omega|=2^3$. Відповідно, подія A складатиметься з $|A|=2^2$ елементарних подій. Тоді*

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

Аналогічно

$$P(B)=\frac{3}{2^3}=\frac{C_3^2}{2^3}=\frac{C_3^1}{2^3}=\frac{3}{8}\ ,$$

$$P(C)=1-\frac{1}{2^3}=\frac{7}{8}\ .$$

Задача 6. *Нехай симетрична монета підкидається доти, доки не випаде герб. Знайти ймовірність того, що ця подія відбудеться при п'ятому підкиданні, за парну кількість підкидань, за непарну кількість підкидань.*

Розв'язання. *Тут елементарним подіями слід вважати такі, що герб випадає при першому, другому, ..., n-ому підкиданні тощо $\Omega=\{\omega_1,\omega_2,...,\omega_n,...\}$. Тут цікаво, що різні елементарні події не є рівноправними. Проте, кожну елементарну ймовірність можна знайти, розв'язавши окрему задачу і застосувавши класичне означення ймовірностей справедливе для рівноправних подій.*

Знайдемо, наприклад, елементарну ймовірність $P(\omega_1)$. *Простір елементарних подій для випадкового експерименту з одним підкиданням буде складатись з двох рівноправних подій* $\Omega_1 = \{\omega_Г, \omega_Р\}$. *Лише одна з них відповідає випаденню герба. Отже,* $P(\omega_1) = \dfrac{1}{2}$.

Для двох підкидань простір елементарних подій складатиметься вже з чотирьох рівноправних елементарних подій $\Omega_2 = \{\omega_{ГГ}, \omega_{РР}, \omega_{ГР}, \omega_{РГ}\}$. *Лише одна з них, що відповідає випаденню решки при першому підкиданні і випаденню герба при другому підкиданні. Отже,* $P(\omega_2) = \dfrac{1}{2^2}$. *Цей процес знаходження інших елементарних ймовірностей можна продовжити далі.*

Нехай герб випав при n*-му підкиданні і нас цікавить саме ця подія. Елементарними подіями вважатимемо послідовність результатів окремих підкидань включно з* n*-м. Простір елементарних подій тут міститиме* 2^n *рівноправних елементарних подій. Нас цікавить лише одна елементарна подія, де монета випаде гербом лише при останньому підкиданні. Отже,* $P(\omega_n) = (1/2)^n$. *При цьому, очевидно, виконується умова нормування* $\sum_{n=1}^{\infty} P(\omega_n) = \sum_{n=1}^{\infty} (1/2)^n = 1$.

Таким чином, ймовірність того, що це відбудеться при п'ятому підкиданні є $P(\omega_5) = (1/2)^5 = 1/32$.

Те, що герб випаде при парному числі підкидань, означає, що якась серія парного числа підкидань буде вдалою для випадення гербу при останньому підкиданні, тобто нас цікавить наступна сума елементарних подій $\sum_{n=1}^{\infty} \omega_{2n}$. *Оскільки всі елементарні події несумісні за означенням, то для ймовірності суми таких елементарних подій, відповідно до властивості ймовірності суми несумісних подій, матимемо*

$$\sum_{n=1}^{\infty} P(\omega_{2n}) = \sum_{n=1}^{\infty} (1/2)^{2n} = 1/3,$$

за непарне число підкидань

$$\sum_{n=1}^{\infty} P(\omega_{2n-1}) = \sum_{n=1}^{\infty} (1/2)^{2n-1} = 2/3.$$

1.5.Геометричне означення ймовірності

Оскільки ми ототожнюємо події з множинами в сенсі теорії множин, то для обчислення ймовірностей цих подій нам необхідно використовувати числові характеристики множин, для можливості їх порівняння. Для скінчених множин з рівноправних елементів такою зручною числовою характеристикою є кількість елементів множини. Така характеристика дозволяє легко порівнювати множини між собою. У разі нерівноправних елементів скінчених множин кількість елементів множин мало про що говорить. Але тут можна ставити у відповідність число не цілій множині, а кожному окремому елементу цієї множини, тобто ввести у розгляд функцію дискретного аргументу. Такими зручними числами є елементарні ймовірності, які можуть бути різними для різних елементів множини. В якості числової характеристики множини в цілому можна взяти суму елементарних ймовірностей, що відповідають кожному її елементу. Такий підхід легко поширюється на злічені множини. У цьому разі немає сенсу оперувати такою характеристикою, як кількість елементів, через їх нескінчену кількість. Проте залишається абсолютно ефективною така характеристика кожного елементу множини, як елементарна ймовірність. При цьому виникає необхідність підсумовувати нескінчену, але злічену кількість доданків (дивись останню задачу). Таким чином ми знову отримуємо можливість характеризувати числом злічену множину.

Для нескінчених множин, елементи яких утворюють неперервну множину ситуація ускладнюється ще більше, порівняно з дискретними зчисленими множинами. Для таких множин вже і елементарні ймовірності втрачають сенс. Проте, аналогічно випадку дискретних множин, кожному елементу множини (кожній точці) доцільно поставити у відповідність число, яке тепер вже не є елементарною ймовірністю. Сукупність всіх таких чисел є нічим іншим, як функцією неперервного аргументу. Серед нескінченої кількості можливих функцій нам надалі знадобляться такі, що в силу певних накладених на них обмежень пов'язані з ймовірностями. Тоді множини можна порівнювати, порівнюючи інтеграли від цих функцій. Самі функції називаються густинами розподілу ймовірностей. Ситуація різко спрощується, якщо всі елементи (точки) неперервних множин є рівноправними (у цьому разі густина розподілу ймовірності є сталою величиною). Тоді множини однієї розмірності знову легко порівнювати. Використовуючи такі їх числові характеристики як довжини – у разі одновимірних множин, площі – у разі двовимірних множин, об'єми – у разі тривимірних множин тощо. Позначаючи зазначені числові міри неперервних множин через $m\left(A\right)$, хоча також вживатимемо і позначення $|A|$, і надаючи йому універсального змісту, матимемо наступний аналог класичного значення ймовірності

$$P\left(A\right)=\frac{m(A)}{m(\Omega)},$$

або

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

Задача 4. *Випадкова точка* A *має рівномірний розподіл у квадраті зі стороною 1. Знайти ймовірності наступної події:* $A = \{$*відстань від точки* A *до фіксованої сторони квадрата не перевищує* $x\}$.

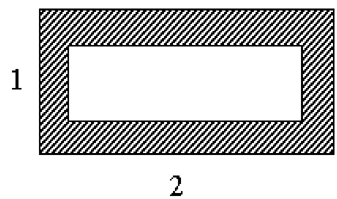
Розв'язання. *Нехай відстань від фіксованої сторони квадрату дорівнює* x. *Тоді елементарною подією можна вважати будь-яке значення цієї відстані таке, що* $0 \le x \le 1$. *Простір елементарних подій утворює сукупність всіх таких значень. Мірою цієї множини, через рівноправність всіх значень, є максимальна довжина цієї відстані* $|\Omega| = 1$. *Події* A, *що відповідає відстані* x *від фіксованої сторони відповідає сукупність точок відрізку* $[0, x]$. *Мірою цієї множини є довжина відповідного відрізку* $|A| = x$. *Отже,* $P(A) = |A| / |\Omega| = x$.

Задача 5. *Випадкова точка* A *має рівномірний розподіл у прямокутнику з сторонами 1 і 2. Знайти ймовірності наступної події:* $A = \{$*відстань від* A *до найближчої сторони прямокутника не перевищує* $x\}$.

Розв'язання.

А. Елементарною подією вважатимемо довільну точку прямокутника. Сукупність всіх точок, що утворюють прямокутник, візьмемо в якості простору елементарних подій. Мірою цієї множини, через рівноправність елементарних подій, є площа прямокутника $|\Omega| = 2$. *Будь-яка сторона прямокутника може бути найближчою до точки* A. *У цьому сенсі всі сторони прямокутника рівноправні. Точки, кожна з яких може найближчою до відповідної сторони утворюють всередині прямокутника полосу вздовж його сторін. Оскільки відстань до найближчої сторони позначено через* x, *то це і буде шириною цієї полоси. Її площа визначить міру множини, що відповідає події – відстань до найближчої сторони не перевищує* x.

Очевидно, що $|A| = 2 - (2-2x)(1-2x) = 2x(3-2x)$. *Відповідно,* $P(A) = x(3-2x)$, $0 \le x \le 1/2$, $P(A) = 1$, $x \ge 1/2$.



Мал. 1.

***Задача 6**. (Бюфона). Площина покреслена паралельними прямими, відстань між якими дорівнює* $2a$. *На площину навмання кинута голка довжини* $2l\,(l < a)$. *Знайти ймовірність того, що голка перетне довільну пряму: 1) голка кинута перпендикулярно до ліній, 2) голка кинута до ліній під кутом* φ, *3) голка кинута під довільним кутом.*

***Розв`язання.** 1) Оскільки нас цікавить лише взаємне розташування голки і найближчої прямої, то в якості простору елементарних подій* Ω *у разі, коли голка кидається перпендикулярно до лінії, слід обрати відрізок прямої довжиною* a, *перпендикулярний лінії. Голка перетинає лінію, якщо центр голки належить частині цього відрізку довжиною* l. *Всі точки цього останнього відрізку і утворюють подію, що нас цікавить. Отже,*

$$P(A) = \frac{|A|}{|\Omega|} = \frac{l}{a}.$$

2) Якщо голка кинута до лінії під кутом φ, *то простір елементарних подій буде той самий, що і у попередньому випадку, а в якості події слід взяти сукупність всіх точок* *частини цього відрізку, довжина якого дорівнює довжині проекції голки на перпендикуляр до лінії. Отже,*

$$P(A) = \frac{|A|}{|\Omega|} = \frac{l\sin(\varphi)}{a}.$$

3) Нехай спочатку голка кидається так, що кут φ змінюється в межах від φ до $\varphi + d\varphi$. Тоді простором елементарних подій буде вже не відрізок довжиною a, перпендикулярний лінії, а прямокутник висотою a і шириною $d\varphi$, тобто

$$|d\Omega| = a\,d\varphi.$$

Шуканою подією також буде прямокутник висотою $l\sin(\varphi)$ і шириною $d\varphi$, тобто $|dA| = l\sin(\varphi)\,d\varphi$.Тепер ймовірність голкою перетнути лінію буде такою

$$P(dA) = \frac{|dA|}{|d\Omega|} = \frac{l\sin(\varphi)d\varphi}{ad\varphi} = \frac{l\sin(\varphi)}{a},$$

як і у попередньому випадку.

Якщо кут φ довільний, тобто змінюється в межах від 0 до π, то простором елементарних подій вже буде інтеграл

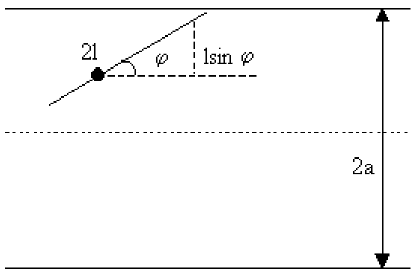
$$|\Omega| = \int_0^{\pi} ad\varphi = a\pi.$$

Подія, що нас цікавить, також визначатиметься інтегралом

$$|A| = \int_0^{\pi} l\sin(\varphi)d\varphi = 2l.$$

Нехай r – відстань від центру голки до найближчої прямої, а φ – кут між голкою і цією прямою. Пара чисел (r,φ) задає положення голки з точністю до вибору конкретної прямої. Оскільки нас цікавить лише взаємне розташування голки і найближчої прямої, то в якості простору елементарних подій

Ω слід обрати прямокутник $\Omega = \{(r,\varphi) : 0 \le r \le a,\ 0 \le \varphi \le \pi\}$. Оскільки всі точки цієї множини рівноправні, то мірою цієї множини є площа відповідного прямокутника $|\Omega| = \pi a$. Таким чином кидання голки на площину рівносильне киданню точки на цей прямокутник. При цьому голка перетинається з прямою тоді і лише тоді, коли виконується нерівність $r \le l\sin(\varphi)$. Іншими словами, коли центр голки є ближчим до лінії, ніж довжина проекції половини голки на напрям, перпендикулярний лінії. Тобто, якщо голка перетинається з прямою, то точка, що їй відповідає, потрапляє всередину фігури, обмеженою кривою $f(\varphi) = l\sin(\varphi)$ та віссю 0φ.



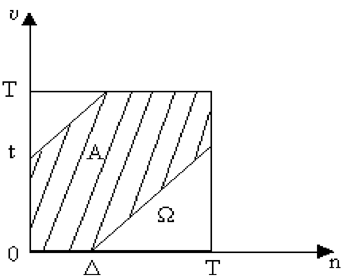
***Мал. 2**.*

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2l}{\pi a}.$$

До речі, експеримент з киданням голки можна використати для експериментального знаходження числа π. Для цього голка кидається багато разів. Частота перетину голкою лінії береться в якості ймовірності цього перетину. Чим ближче частота до ймовірності,тим точнішим є експериментальне значення числа π.

***Задача** 7. За проміжок часу $\left[0,T\right]$ у випадковий момент n з'являється сигнал довжиною Δ. Приймач вмикається у випадковий момент $\upsilon \in \left[0,T\right]$ на час t. Вважаючи, що точка $\left(n,\upsilon\right)$ рівномірно розподілена в квадраті $\left[0,T\right]\times\left[0,T\right]$, знайти ймовірність виявлення сигналу.*

***Розв`язання.** У якості елементарної події слід взяти довільну точку $\left(n,\upsilon\right)$, а в якості простору елементарних подій – квадрат $\Omega = \left\{\left(n,\upsilon\right): 0 \le n \le T, 0 \le \upsilon \le T\right\}$. Мірою цієї множини є площа відповідного квадрату $\left|\Omega\right| = T^2$. Якщо сигнал з'являється в момент часу $n = 0$, то він може бути виявленим, якщо приймач вмикається у будь – який момент проміжку $\left[0,t\right]$. Якщо сигнал з'являється у момент часу n $\left(0 \le n \le \Delta\right)$, то він може бути виявленим, якщо приймач вмикається у будь – який момент проміжку $\left[0,t+n\right]$. Якщо сигнал з'являється у момент часу*



***Мал. 3**.*

$\left(\Delta < n \le T\right)$, то він може бути виявленим, якщо приймач вмикається у будь – який момент проміжку $\left[n,t+n\right]$. Таким чином, подія A, що нас цікавить, відповідає площі заштрихованої фігури. Її площу можна знайти, якщо від площі

квадрату відняти площі двох не заштрихованих прямокутних трикутників

$$\left|A\right| = T^2 - \frac{1}{2}\left(T - t\right)^2 - \frac{1}{2}\left(T - \Delta\right)^2 .$$

Відповідно,

$$P\left(A\right) = \frac{\left|A\right|}{\left|\Omega\right|} = \frac{T^2 - \frac{1}{2}\left(T - t\right)^2 - \frac{1}{2}\left(T - A\right)^2}{T^2} =$$

$$= 1 - \frac{1}{2}\left(1 - \frac{t}{T}\right)^2 - \frac{1}{2}\left(1 - \frac{\Delta}{T}\right)^2 .$$

1.6.Формула повної ймовірності

Класичне і геометричне означення ймовірності вичерпують собою алгоритми безпосереднього обчислення ймовірностей. В усіх інших випадках або доцільно, або без варіантів необхідне використання складніших алгоритмів обчислення ймовірностей. Типовим для теорії ймовірностей є структуризація випадкового експерименту, зокрема, розбиття його на незалежні або маже незалежні етапи, комбіноване використання теорем додавання і множення ймовірностей тощо. Однією з найпростіших схем такого типу обчислення ймовірностей є формула повної ймовірності. Вона застосовна тоді, коли подія, що нас цікавить, є наслідком реалізації однієї з багатьох взаємовиключних подій-гіпотез, які разом утворюють повну систему гіпотез, тобто реалізація будь-якої з цих гіпотез є достовірною подією. Кожна з гіпотез приводить до короткого (двоетапного) логічного ланцюжка, наслідком реалізації якого і є подія, що нас цікавить. розглянемо цю схему докладніше.

Нехай A довільна подія, а події $B_1, B_2, ..., B_n$ попарно несумісні і такі, що $B_1 + B_2 + ... + B_n = \Omega$. При цьому вважатимемо, що $P(B_k) > 0,\ k = 1, 2, ..., n$ і A має ненульовий перетин з кожною з множин $B_1, B_2, ..., B_n$. Тоді матиме місце формула

$$P(A) = \sum_{k=1}^{n} P(B_k) P(A \mid B_k).$$

Для доведення цієї формули використаємо аксіому адитивності в аксіоматичному означенні ймовірності, для чого попередньо представимо подію A *сумою наступних несумісних подій*

$$A = A \cdot B_1 + A \cdot B_2 + ... + A \cdot B_n.$$

Тоді, на основі згаданої аксіоми,

$$P(A) = P(A \cdot B_1) + P(A \cdot B_2) + ... + P(A \cdot B_n).$$

Застосовуючи до кожного доданку у правиці рівняння теорему множення ймовірностей, ми і отримаємо формулу повної ймовірності.

Важливим наслідком з формули повної ймовірності є формули Баєса. Для їх отримання скористаємось тотожністю

$$P(A) P(B_k \mid A) = P(B_k) P(A \mid B_k),$$

яка випливає з комутативності добутку двох подій і відповідних їм виразів для ймовірності

$$P(A \cdot B_k) = P(B_k) P(A \mid B_k),$$

$$P(A \cdot B_k) = P(A) P(B_k \mid A).$$

З зазначеної тотожності отримуємо формулу Баєса

$$P(B_k \mid A) = \frac{P(B_k) P(A \mid B_k)}{P(A)}.$$

Ця формула дозволяє знайти ймовірність реалізації гіпотези B_k *- функцію* $P(B_k \mid A)$ *при умові, що подія* A *відбулась.*

Розглянемо декілька задач на використання формули повної ймовірності.

Задача 1. *Серед 25 екзаменаційних білетів 5 таких, що студенти їх знають. Два студенти по черзі беруть по одному білету. Знайти ймовірність того, що*

А) перший студент взяв такий білет, що знає,

В) другий студент взяв такий білет, що знає,

С) обидва студенти взяли такі білети, що знають.

Розв'язання. *Нехай подія A – перший студент взяв білет, що знає, подія B – другий студент взяв білет, що знає, подія С- обидва студенти взяли білети, що знають. Тоді для першого студента простір елементарних подій складається з 25 елементарних подій $|\Omega| = 25$, подія A складається з 5 елементарних подій $|A| = 5$ і за класичних означенням ймовірності ймовірність того, що перший студент взяв білет, що знає,*

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{5}{25} = \frac{1}{5}.$$

Для другого студента простір елементарних подій складається з 24 елементарних подій $|\Omega| = 24$, оскільки після того, як перший студент витягнув один білет їх загальна кількість зменшилась на один. Подія B складається або з 5 елементарних подій, якщо перший студент взяв поганий білет, і

$$P\left(B / \overline{A}\right) = \frac{|B / \overline{A}|}{|\Omega|} = \frac{5}{25} = \frac{1}{5} ,$$

або з 4 елементарних подій, якщо перший студент взяв білет, що знає, і

$$\mathrm{P}\left(B / A\right) = \frac{|B / A|}{|\Omega|} = \frac{4}{24} = \frac{1}{6} .$$

Оскільки можливі обидва варіанти, то для визначення шуканої ймовірності слід скористатись формулою повної ймовірності. Тоді

$$P(B) = P(A)\cdot P(B/A) + P(A)P(B/A) = \frac{1}{5}\cdot\frac{1}{6} + \frac{4}{5}\cdot\frac{5}{24} = \frac{1}{5}\ .$$

Ймовірність того, що обидва студенти взяли білет, що знають, визначається відповідно до теореми множення ймовірностей

$$P(A\cdot B) = P(A)P(B/A) = \frac{1}{5}\cdot\frac{1}{6} = \frac{1}{30}\ .$$

***Задача 8**. При переливанні крові слід враховувати групу крові донора і хворого. Людині, що має четверту групу крові, можна переливати кров довільної групи; людині з другою або третьою групою крові можна переливати кров або тої ж групи, або першої; людині з першою групою крові можна переливати лише кров першої групи. Серед населення 33.7 % мають першу, 37.5 % – другу, 20.9 третю і 7.9 % – четверту групи крові.*

Знайти ймовірність того, що випадково взятому хворому можна перелити кров випадково взятого донора.

***Розв`язання.** Нехай B_1 – подія, відповідно до якої донор має першу групу крові; B_2 – другу; B_3 – третю; B_4 – четверту. Якщо подія A полягає у тому, що у висліді переливання крові хворий виживе, то ймовірність цієї події визначається формулою повної ймовірності*

$$P(A) = \sum_{n=1}^{4} P(B_n)P(A/B_n).$$

Очевидно, $P(B_1) = 0.337$, $P(B_2) = 0.375$, $P(B_3) = 0.209$, $P(B_4) = 0.079$, $P(A/B_1) = 0.337$, *оскільки хворий повинен мати виключно першу групу крові;* $P(A/B_2) = 0.337 + 0.375 = 0.712$ *– оскільки хворий повинен мати або першу, або другу групи крові;*

аналогічно $P\left(A/B_3\right)=0.337+0.209=0.546$*;* $P\left(A/B_4\right)=1$, *оскільки хворий у цьому разі може мати довільну групу крові. У результаті*

$$P\left(A\right)=0.337\cdot 0.337+0.375\cdot 0.712+0.209\cdot 0.546+0.079\cdot 1=0.57$$

1.7.Послідовність незалежних випробувань

Інколи випадковий експеримент можна представити послідовністю простіших випадкових експериментів, що відбуваються в одних і тих же умовах і результати яких не пов'язані між собою. Наприклад, розглянемо випадковий експеримент, що полягає в десятиразовому підкиданні монети. Ймовірність довільного наслідку такого експерименту для симетричної монети можна обчислити, використовуючи класичне означення ймовірності, застосовуючи його відразу до всієї серії підкидань. Тут елементарною подією можна було б вважати довільний результат всієї серії підкидань. Таких елементарних подій у нас буде 2^{10}. Інакше, можна класичне означення ймовірності застосувати лише до одного підкидання. Тоді елементарних подій буде лише дві. Кожне підкидання у цьому разі можна розглядати як незалежне випробування. Будь-яку подію, пов'язану з усією серією підкидань, можна розглянути як добуток незалежних подій, пов'язаних з окремими підкиданнями. При цьому, підраховувати доведеться вже не елементарні події, яких у першому підході 2^{10}, а незалежні випробування, тобто окремі підкидання, що зробити значно легше. При будь-якій серії підкидань процедура підрахунку випробувань, в яких спостерігається потрібний наслідок, буде стандартною. Її можна виконати один раз і отримати універсальний результат, вірний для довільної серії підкидань. Якщо в незалежних випробуваннях даної серії може бути лише два наслідки, які зручно назвати успіхом і невдачею, тобто простір елементарних подій для окремого випробування складається лише з двох елементарних подій, то відповідь на питання, якою буде ймовірність m успіхів в серії з n незалежних випробувань визначатиметься формулою Бернуллі, а відповідна схема обчислення ймовірності - схемою Бернуллі,

$$P_n(m) = C_n^m p^m (1-p)^{n-m} .$$

Тут p - ймовірність успіху в окремому випробуванні - ймовірність однієї з двох елементарних подій простору елементарних подій окремого випробування, яке ми умовно називаємо успіхом. Ймовірність другої елементарної події, яка умовно називається невдачею, $1-p$. Ця формула є очевидним наслідком того, що m успіхів можна C_n^m способами розмістити серед послідовності в n випробувань. Насправді схема Бернуллі тільки у разі $p = 1/2$ в принциповому сенсі еквівалентна класичному означенню ймовірності. Елементарні ймовірності в окремому випробуванні можуть, насправді, бути довільними. Ми взагалі можемо вважати їх параметрами теорії, як це і є в формулі Бернуллі. Тобто формула Бернуллі може виходити за рамки можливостей класичного означення ймовірностей.

Розглянемо схему Бернуллі детальніше. Зробимо це на прикладі підкидання монети. Монета може бути не обов’язково симетрична. Позначимо ймовірність випадення герба (успіх) при одному підкиданні p, а випадення герба (невдача) $q = 1-p$. Якщо монета підкидається n разів, то простір елементарних подій складатиметься з елементарних подій, кожна з яких є результатом всієї серії підкидань. Наприклад, якщо герб не випав жодного разу, то відповідну елементарну подію можна позначити $\omega_{\text{Р...Р}}$. Якщо герб випав лише перший раз, то $\omega_{\text{ГР...Р}}$ тощо. Отже,

$$\Omega = \left\{ \omega_{\text{Р}\ldots\text{Р}}, \omega_{\text{ГР}\ldots\text{Р}}, \omega_{\text{РГ}\ldots\text{Р}}, \ldots \right\}.$$

У такому просторі вже не всі елементарні події будуть рівноправні. Рівноправними будуть лише елементарні події, що відповідають фіксованій кількості випадень герба. Наприклад, всі елементарні події, що відповідають випаденню одного герба $A_n(1)$ будуть рівноправними. Всі елементарні події, що відповідають випаденню двох гербів $A_n(2)$ тощо. Елементарні події, що входять в різні події $A_n(0), A_n(1)$, $A_n(2)$, …, $A_n(n)$ рівноправними не будуть. Сума всіх таких подій дорівнюватиме простору елементарних подій

$$\Omega = \sum_{m=0}^{n} A_n(m).$$

Оскільки ймовірність суми несумісних подій дорівнює сумі їх ймовірностей, то маємо умову нормування

$$P(\Omega) = P\left(\sum_{m=0}^{n} A_n(m) \right) = \sum_{m=0}^{n} P\big(A_n(m)\big) = \sum_{m=0}^{n} P_n(m) = 1.$$

Тут $P_n(m)$ - ймовірність випадення m гербів при n підкиданнях.

Для знаходження ймовірностей $P_n(m)$ Можна застосувати класичне означення ймовірності, оскільки всі елементарні події, що входять у відповідний простір елементарних подій будуть рівноправними. Їх кількість визначатиметься кількістю способів, якими можна розташувати m гербів на n місцях з урахуванням того факту, що перестановки гербів між собою відповідають одній і тій самій елементарній події, тобто кількістю сполучень з n елементів по m, тобто C_n^m. Таким чином $P_n(m) = C_n^m P\left(\omega_{\text{Г}_1\ldots\text{Г}_m\text{Р}_1,\ldots\text{Р}_{n-m}}\right)$. Для знаходження ймовірності довільної елементарної події, зокрема наведеної вище, слід врахувати, що елементарну подію для всього складного випадкового експерименту можна представити добутком елементарних подій для окремих незалежних його складових $\omega_{\Gamma_1 ... \Gamma_m \mathrm{P}_1, ... \mathrm{P}_{n-m}} = \omega_{\Gamma_1 .} ... \omega_{\Gamma_m} \omega_{\mathrm{P}_1} ... \omega_{\mathrm{P}_{n-m}}$. Оскільки кожна з елементарних подій добутку є незалежними, оскільки стосуються різних простих випадкових експериментів, з яких складається весь випадковий експеримент, то ймовірність такого добутку дорівнюватиме добутку ймовірностей

$$P(\omega_{\Gamma_1 ... \Gamma_m \mathrm{P}_1, ... \mathrm{P}_{n-m}}) = P(\omega_{\Gamma_1 .}) ... P(\omega_{\Gamma_m}) P(\omega_{\mathrm{P}_1}) ... P(\omega_{\mathrm{P}_{n-m}}) = p^m q^{n-m} .$$

У висліді ми і отримуємо формулу Бернулі

$$P_n(m) = C_n^m p^m (1-p)^{n-m} .$$

Подальшим узагальненням формули Бернуллі є поліноміальна схема. Вона виникає, якщо кількість наслідків одного випробування послідовності незалежних випробувань більше за два, наприклад, N. Тоді ймовірність того, що у випадку послідовності n випробувань буде m_1 раз зустрічатись наслідок 1, m_2 раз наслідок 2, ..., m_N раз наслідок N при тому, що $m_1 + m_2 + \cdots + m_N = n$, матиме вигляд

$$P_n(m_1, m_2, ..., m_N) = \frac{n!}{m_1 ! m_2 ! \cdots m_N} p_1^{m_1} p_2^{m_2} \cdots p_N^{m_N} .$$

Для формули Бернуллі існує декілька наближених асимптотичних формул, які спрощують обчислення ймовірностей у разі великих n.

Одна з таких асимптотичних формул виникає, якщо $n \to \infty$ але при цьому $p \to 0$ так, що $np = \lambda = const$. Тоді

$$P_n(m) \approx \frac{\lambda^m}{m!} \exp(-\lambda) .$$

Остання формула називається формулою Пуассона. Із зростанням λ формула Пуасона прямує до нормального розподілу з математичним очікуванням λ і дисперсією λ.

Якщо нас цікавить ймовірність приналежності кількості успіхів певному інтервалу, то слід виконати наступне підсумовування

$$P_n(m_1 \le m \le m_2) = \sum_{m=m_1}^{m_2} P_n(m).$$

Задача 1. *Довести, що для* $n \to \infty$ *і* $p \to 0$ *так, що* $n \cdot p \to \lambda$, $0 < \lambda < \infty$, $P_n(m) = \frac{\lambda^m}{m!}\exp(-\lambda)$.

Розв`язання. *Покладемо* $n \cdot p = \lambda$ *і представимо ймовірність* $P_n(m)$ *у вигляді*

$$P_n(m) = C_n^m p^m (1-p)^{n-m} = \frac{n(n-1)\dots(n-m+1)}{m!}\left(\frac{\lambda_m}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-m} =$$

$$= \frac{\lambda_n^m}{m!}\left(1-\frac{1}{n}\right)\dots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^n\left(1-\frac{\lambda_n}{n}\right)^{-m}.$$

Перейдемо тепер до границі $n \to \infty$

$$\lim_{n\to\infty} P_n(m) = \lim_{n\to\infty}\left[\frac{\lambda_n^m}{m!}\left(1-\frac{1}{n}\right)\dots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^n\left(1-\frac{\lambda_n}{n}\right)^{-m}\right] =$$

$$= \frac{\lambda_n^m}{m!}\lim_{n\to\infty}\left[\left(1-\frac{1}{n}\right)\dots\left(1-\frac{m-1}{n}\right)\right]\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^n \lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{-m} =$$

$$= \frac{\lambda^m}{m!}\exp(-\lambda).$$

Тут ми використали той факт, що

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^n = \exp(-\lambda),$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{-m}=1,$$

$$\lim_{n\to\infty}\left[\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{m-1}{n}\right)\right]=1.$$

Інша асимптотична формула виникає, якщо $n\to\infty$ але при цьому p довільне, тобто умова $np=\lambda=const$ відсутня. Тоді має місце локальна формула Муавра-Лапласа

$$P_n(m)\approx\frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-\frac{(m-np)^2}{2np(1-p)}\right).$$

Ця формула має високу точність для значень m, що лежать в околі точки np, точніше в інтервалі $(np-\sqrt{np(1-p)}, np+\sqrt{np(1-p)})$. Цю формулу можна записати і в іншому вигляді

$$P_n(m)\approx\frac{1}{\sqrt{np(1-p)}}f(x_m),$$

де

$$f(x_m)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x_m^2}{2}\right)$$

так звана густина нормального розподілу випадкової величини з нульовим математичним очікуванням і одиничною дисперсією, що має наступну структуру

$$x_m=\frac{m-np}{\sqrt{npq}}.$$

Якщо нас цікавить приналежність значень даної випадкової величини певному інтервалу $(a\le x_m\le b)$, то за тих же умов має місце інтегральна формула Муавра-Лапласа

$$P(a\le x_m\le b)\approx\frac{1}{\sqrt{2\pi}}\int_a^b\exp\left(-\frac{x^2}{2}\right)dx.$$

Останній результат можна виразити через функцію Лапласа

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{y^2}{2}\right) dy$$

так

$$P(a \le x \le b) = \Phi_0(b) - \Phi_0(a) ,$$

або через функцію нормального розподілу випадкової величини x у подібний спосіб

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(y^2 / 2\right) dy ,$$

а саме

$$P(a \le x \le b) = F(b) - F(a) .$$

Очевидно, функція Лапласа є непарною

$$\Phi_0(-x) = \frac{1}{\sqrt{2\pi}} \int_0^{-x} \exp\left(-\frac{y^2}{2}\right) dy =$$

$$= -\frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{y^2}{2}\right) dy = -\Phi_0(x) ,$$

де ми виконали заміну змінної інтегрування $y \to -y$.

Якщо записати результат інтегральної теореми Муавра-Лапласа безпосередньо для кількості успіхів, то матимемо

$$P(m_1 \le m \le m_2) = \Phi_0\left(\frac{m_2 - np}{\sqrt{np(1-p)}}\right) - \Phi_0\left(\frac{m_1 - np}{\sqrt{np(1-p)}}\right)$$

Останній результат можна отримати і використовуючи лише локальну теорему Муавра-Лапласа. Дійсно, якщо в схемі Бернулі нас цікавить ймовірність приналежності кількості успіхів до певного інтервалу значень $m_1 \le m \le m_2$, то можна використати локальну теорему Муавра-Лапласа і просто знайти суму всіх ймовірностей успіху для $m_1 \le m \le m_2$

$$P_n(m_1 \le m \le m_2) = \sum_{m=m_1}^{m_2} P_n(m).$$

Тут ймовірності під знаком суми визначаються або формулою Бернулі, якщо підсумовування технічно можливе при не дуже великих кількостях випробувань та успіхів, або формулою Пуасона, якщо підсумовування з використанням формули Бернулі технічно неможливе, але виконуються умови застосовності формули Пуасона, або нормального розподілу відповідно до умов виконання локальної теореми Муавра-Лапласа. В останньому разі можна перейти від підсумовування до інтегрування за m, вважаючи її неперервною змінною. Тоді

$$P_n(m_1 \le m \le m_2) = \frac{1}{\sqrt{2\pi np(1-p)}} \int_{m_1}^{m_2} \exp\left(-\frac{(m-np)^2}{2np(1-p)} \right) dm.$$

Ясно, що інтегрування дасть не такий точний результат, як підсумовування, оскільки містить додаткове припущення, що інтегранда мало змінюється при зміні індексу підсумовування на одиницю.

Обчислення останнього інтегралу можна привести до обчислення функції Лапласа, яка обговорювалась вище. Для цього здійснимо заміну змінних під знаком інтегралу

$$x = \frac{m-np}{\sqrt{np(1-p)}}.$$

Тепер інтеграл матиме вигляд

$$P_n(m_1 \le m \le m_2) = \frac{1}{\sqrt{2\pi}} \int_{\frac{m_1-np}{\sqrt{np(1-p)}}}^{\frac{m_2-np}{\sqrt{np(1-p)}}} \exp\left(-\frac{x^2}{2} \right) dx.$$

Останній результат можна записати як

$$P_n(m_1 \le m \le m_2) = \Phi_0\left(\frac{m_2-np}{\sqrt{np(1-p)}} \right) - \Phi_0\left(\frac{m_1-np}{\sqrt{np(1-p)}} \right).$$

Очевидно, що цей результат збігається з результатом інтегральної теореми Муавра-Лапласа.

Великого практичного значення цим формулам надавали в до комп'ютерну епоху, коли обчислення величини C_n^m для великих m і n становило певну проблему.

З теоретичної точки зору ці асимптотичні формули є одним з варіантів центральної межової теореми, оскільки показують, що функція розподілу випадкової величини, відмінна від нормальної, за певних умов прямує саме до неї.

Інше важливе прикладне значення інтегральної теореми Муавра-Лапласа полягає у тому, що вона дозволяє оцінити близькість частоти випадкової події до її ймовірності в залежності від кількості випробувань. Відповідно до формули Мізеса

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

ми лише констатуємо остаточний результат недосяжної експериментально операції межового переходу. А як оцінити точність цієї рівності при скінченій кількості випробувань? Для цього розглянемо n випробувань, у висліді яких m разів спостерігалась потрібна нам подія. Очевидно, частотою цієї події буде m/n, а її ймовірність позначимо через p. Нехай величина Δ характеризує відхилення частоти від ймовірності. Знайдемо ймовірність того, що відхилення частоти від математичного очікування не перевищує Δ, тобто $P\left(\left|m/n-p\right|<\Delta\right)$. Структура випадкової величини $\left|m/n-p\right|$не дозволяє застосувати безпосередньо до неї теорему Муавра-Лапласа. Але візьмемо до уваги, що випадкова величина $x_m = \dfrac{m-np}{\sqrt{np(1-p)}}$, згідно з інтегральною теоремою Муавра-Лапласа, буде розподілена за нормальним законом з нульовим математичним очікуванням і одиничною дисперсією, тобто незалежними від параметрів випадкової величини. Спробуємо перейти до цієї останньої випадкової величини виходячи з випадкової величини, заданої нам за умовою

$$P\left(\left|\frac{m}{n}-p\right|<\Delta\right)=P\left(-\Delta<\frac{m}{n}-p<\Delta\right)=$$

$$=P\left(-\Delta\sqrt{n}<\frac{m-np}{\sqrt{n}}<\Delta\sqrt{n}\right)=$$

$$=P\left(-\Delta\sqrt{\frac{n}{p(1-p)}}<\frac{m-np}{\sqrt{np(1-p)}}<\Delta\sqrt{\frac{n}{p(1-p)}}\right).$$

Останню ймовірність вже можна оцінити, використовуючи інтегральну теорему Муавра-Лапласа

$$P\left(-\Delta\sqrt{\frac{n}{p(1-p)}}<\frac{m-np}{\sqrt{np(1-p)}}<\Delta\sqrt{\frac{n}{p(1-p)}}\right)=$$

$$=\Phi_0\left(\Delta\sqrt{\frac{n}{p(1-p)}}\right)-\Phi_0\left(-\Delta\sqrt{\frac{n}{p(1-p)}}\right)=$$

$$=2\Phi_0\left(\Delta\sqrt{\frac{n}{p(1-p)}}\right).$$

Отже,

$$P\left(\left|\frac{m}{n}-p\right|<\Delta\right)=2\Phi_0\left(\Delta\sqrt{\frac{n}{p(1-p)}}\right).$$

Остання формула є зручною, якщо ймовірність p відома. Тоді отриманий результат для ймовірності є точним. Якщо ж ймовірність p невідома, то замість точного результату ми можемо послугуватись лише оцінкою. Її легко отримати з тих міркувань, що $p(1-p)\le 1/4$. Тоді $1/p(1-p)\ge 4$, а $\Phi_0\left(\Delta\sqrt{\dfrac{n}{p(1-p)}}\right)\ge\Phi_0\left(2\Delta\sqrt{n}\right)$. Шукана оцінка матиме вигляд

$$P\left(\left|\frac{m}{n}-p\right|<\Delta\right)\ge 2\Phi_0\left(2\Delta\sqrt{n}\right).$$

Задача 2*. Вважаючи $p=1/2$, визначити, скільки підкидань монети слід здійснити, щоб із ймовірністю не меншою, ніж 0.9 , відносна частота m/n відрізнялась від 1/2 на 0.05.*

Розв`язання. *Скористаємось нерівністю*

$$\mathrm{P}\left(\left|\frac{m}{n}-p\right|<\Delta\right)\ge 2\Phi_0\left(2\Delta\sqrt{n}\right)=1-2\alpha .$$

Тут $1-2\alpha=0.9$, $\Delta=0.05$, $p=1/2$, m/n – зазначена в умові частота, n – шукана кількість підкидань. Останню знайдемо з умови

$$2\Phi_0\left(0.1\sqrt{n}\right)=0.9 .$$

За таблицею для функції $\Phi_0\left(x\right)$ знайдемо, що умова виконується, якщо аргумент цієї функції дорівнює 1.6. Отже, $0.1\sqrt{n}=1.6$. Звідси $n=256$.

Задача 3*. Нехай N – досить велика кількість однотипних виробів, з яких M – браковані. Тоді природно вважати, що навмання взятий виріб є бракованим з ймовірністю $p=M/N$. Ця ймовірність існує, проте вона може бути невідомою. Для визначення цієї ймовірності на практиці проводять вибірковий контроль: навмання вибирають n виробів і за відносною частотою m/n оцінюють ймовірність p (m – кількість бракованих виробів серед n взятих). Нехай на основі вибірки у 10000 виробів знайдена відповідна частота $m/n=0.0325$.*

Оцінити ймовірність того, що знайдена частота відрізняється від ймовірності менше ніж на $\Delta = 0.005$.

Розв`язання. *Використаємо нерівність*

$$P\left(\left|\frac{m}{n} - p\right| < \Delta\right) \geq 2\Phi_0\left(2\Delta\sqrt{n}\right).$$

За таблицею для функції $\Phi_0\left(x\right)$ *знайдемо, що* $\Phi_0(2 \cdot 0.005 \cdot \sqrt{10000}) = \Phi_0(1) = 0.3413$.

Отже,

$$P\left(\left|\frac{m}{n} - p\right| < 0.005\right) \geq 0.683 .$$

1.8.Задачі для самостійного розв`язання

Варіант 1.

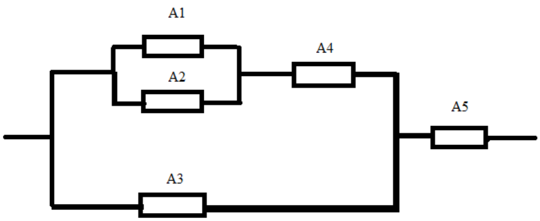
1.В колоді 36 карт. Знайти ймовірності подій:

А - чотири шістки розташовані поряд;

В - місця розташування шісток утворюють арифметичну прогресію з кроком 6.

Побудувати простір елементарних подій і події А і В.

2.Електричне коло складене з елементів A_k , $k = 1, 2, ..., 5$ за схемою, наведеною на малюнку. При виході з ладу довільного елементу коло в місці його ввімкнення розривається. Ймовірність виходу з ладу за даний період елементу A_k дорівнює p_k , $k = 1, 2, ..., 5$. Знайти ймовірність того, що за розглянутий період по колу може проходити електричний струм.



Мал. 4.

3.Під час випробувань було встановлено, що ймовірність безвідмовного спрацьовування реле при відсутності завад дорівнює 0.99, при перегріві - 0.95, при вібрації - 0.9, при вібрації і перегріві - 0.8. Знайти ймовірність відмови цього реле при роботі у спекотній країні, якщо ймовірність перегріву дорівнює 0.2, ймовірність вібрації - 0.1. Вважати перегрів і вібрацію незалежними подіями.

Варіант 2.

1.На полиці у випадковому порядку розташовані 40 книжок, серед яких знаходиться тритомник Т.Г. Шевченка. Знайти ймовірність того, що ці томи розташовані у порядку зростання їх номерів але не обов'язково поряд.

Побудувати простір елементарних подій.

2.Перевезення вантажів для підприємства забезпечують два автогосподарства, які з цією метою щодня в першу зміну мають виділити по одному автомобілю. Ймовірність виходу автомобіля на лінію в першому автогосподарстві дорівнює 0.7, а в другому - 0.6. Знайти ймовірність того, що в першу зміну на підприємстві перевозитимуться вантажі.

3.Під час випробувань було встановлено, що ймовірність безвідмовного спрацьовування реле при відсутності завад дорівнює 0.99, при перегріві - 0.95, при вібрації - 0.9, при вібрації і перегріві - 0.8. Знайти ймовірність відмови цього реле при роботі у пересувній лабораторії, якщо ймовірність перегріву дорівнює 0.1, ймовірність вібрації - 0.3. Вважати перегрів і вібрацію незалежними подіями.

Варіант 3.

1.Підкинуті три монети. Побудувати простір елементарних подій і подій та знайти ймовірності цих подій:

A - третя монета випала гербом догори;

B - випав рівно один герб;

C - випало рівно два герба;

D - випало не більше двох гербів.

2.Прилад складається з трьох вузлів, які працюють незалежно один від одного, причому другий і третій вузли взаємозамінні. Ймовірності виходу з ладу вузлів на заданому часовому проміжку становлять відповідно 0.2, 0.3 і 0.4. Знайти ймовірність того, що протягом заданого часу прилад працюватиме.

3.Є п'ять урн. У першій, другій і третій урнах знаходяться по два білих і три чорних кулі, у четвертій і п'ятій - по одній білій і одній чорній кулі. Випадково обирається урна і з неї витягається куля. Яка ймовірність, що ця куля виявиться білою?

Варіант 4.

1.Підкидається гральний кубик. Побудувати простір елементарних подій і подій та знайти ймовірності цих подій:

A - випало число, що не перевищує 2;

B - випало число, що не перевищує 3.

Знайти суму, різницю, добуток і заперечення цих подій.

2.При масовому виготовленні виробів брак становить у середньому 1.5% загальної кількості всіх виробів. З поміж придатних виробів 85.3% становлять вироби 1-го сорту. Знайти ймовірність того, що навмання взятий виріб належить до першого сорту.

3.Є п'ять урн. У першій, другій і третій урнах знаходяться по два білих і три чорних кулі, у четвертій і п'ятій - по одній білій і одній чорній кулі. Випадково обирається урна і з неї витягається куля. Яка ймовірність, що ця куля виявиться чорною?

Варіант 5.

1.В колоді 52 карт. Знайти ймовірності подій:

A - чотири шістки розташовані поряд;

В - місця розташування шісток утворюють арифметичну прогресію з кроком 5.

2.У цеху є три резервні двигуни, для кожного з яких ймовірність ввімкнення у даний момент дорівнює 0.3. Знайти ймовірність того, що у даний момент ввімкнено:

1) принаймні два двигуни;

2) принаймні один двигун.

3.У першій урні знаходяться 2 білих і 8 чорних кульки, у другій - 1 чорна і 4 білих кульки. З кожної втрачено по одній кульці невідомого кольору. Всі кульки, що залишились, зсипані до третьої урни. Знайти ймовірність того, що з першої і другої урн була втрачена чорна кулька при умові, що з третьої урни витягнули чорну кульку.

Варіант 6.

1.Підкинуті три монети. Побудувати простір елементарних подій і подій та знайти ймовірності цих подій:

А - всі три монети випали гербом догори;

В - перша монета випала гербом догори;

С - випало рівно два герба;

D - випало не більше одного герба.

2.Від аеровокзалу відправились два автобуси. Ймовірність своєчасного прибуття кожного з них дорівнює 0.92. Знайти ймовірність такої події:

А - обидва автобуси прибули своєчасно;

В - обидва автобуси запізнились;

С - тільки один автобус прибув своєчасно.

3.При переливанні крові слід враховувати групу крові донора і хворого. Людині, що має четверту групу крові, можна переливати кров довільної групи; людині з другою або третьою групою крові можна переливати кров або тої ж групи, або першої; людині з першою групою крові можна переливати лише кров першої групи. Серед населення 30% мають першу, 35% – другу, 25 третю і 10% – четверту групи крові. Знайти ймовірність того, що хворому перелили другу крупу крові при умові що він вижив.

Варіант 7.

1.Яка ймовірність того, що чотиризначний номер випадково взятого автомобіля:

A - має всі різні цифри;

B - має лише дві однакові цифри;

C - має дві пари однакових цифр;

D -має лише три однакові цифри;

E - має всі однакові цифри.

2.Партія містить 12 стандартних і чотири нестандартні деталі. Навмання беруть три деталі. Знайти ймовірність того, що серед узятих деталей:

1) усі три нестандартні;

2) принаймні одна стандартна;

3) не менше як дві стандартні.

3.У першій урні знаходяться 3 білих і 7 чорних кульки, у другій - 2 чорна і 4 білих кульки. З кожної втрачено по одній кульці невідомого кольору. Всі кульки, що залишились, зсипані до третьої урни. Знайти ймовірність того, що кулька витягнута з третьої урни, виявиться чорною.

Варіант 8.

1.В записаному телефонному номері стерлись три останні цифри. Знайти ймовірності подій:

A - цифри, що стерлись, різні але відмінні від цифр 1, 2, 3;

B - стерлись однакові цифри;

C - дві цифри, що стерлись, збігаються.

2.Є дві партії деталей. У першій партії сім придатних і три бракованих деталі. У другій - 10 придатних і чотири бракованих. Із кожної партії навмання беруть по одній деталі. Знайти ймовірність події:

1) обидві деталі придатні;

2) обидві деталі браковані;

3) одна деталь придатна, а друга бракована.

3.Є п'ять урн. У першій, другій і третій урнах знаходяться по два білих і три чорних кулі, у четвертій і п'ятій - по одній білій і одній чорній кулі. Випадково обирається урна і з неї витягається куля. Яка ймовірність, що при цьому обрані четверта або п'ята кулі при умові, що витягнута біла куля?

Варіант 9.

1.На кожній з п'яти карток написані літери Т, А, Р, А, С. Яка ймовірність того, що викладаючи їх в ряд навмання вдасться написати ім’я ТАРАС? Побудувати простір елементарних подій і відповідної події.

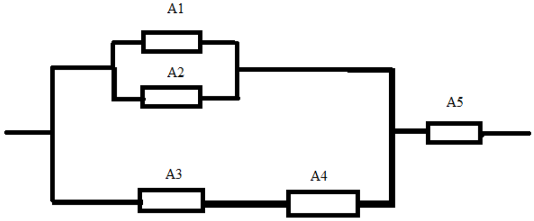
2.Облік щодо використаних запасних частин показав, що в разі ремонту двигуна деталь № 1 замінювалась у середньому в 35% випадків, деталь № 2 - у 30% випадків, а обидві деталі у 28% випадків. Знайти ймовірність того, що у двигуні, який надійшов у ремонт, замінюватиметься деталь № 2 за умови, що деталь № 1 замінена, і навпаки.

3.В будівельному загоні 70% першокурсників і 30% другокурсників. Серед першокурсників дівчата складають 10%, а серед другокурсників - 5%. Всі дівчата по черзі чергують на кухні. Яка ймовірність, що у випадково обраний день на кухні чергує першокурсниця?

Варіант 10.

1.На кожній з чотирнадцяти карток написані літери Е, Л, Е, К, Т, Р, И, Ф, І, К, А, Ц, І, Я. Яка ймовірність того, що викладаючи їх в ряд навмання вдасться написати слово ЕЛЕКТРИФІКАЦІЯ? Побудувати простір елементарних подій і відповідної події.

2.Електричне коло складене з елементів A_k, $k=1,2,...,5$ за схемою, наведеною на малюнку. При виході з ладу довільного елементу коло в місці його ввімкнення розривається. Ймовірність виходу з ладу за даний період елементу A_k дорівнює p_k, $k=1,2,...,5$. Знайти ймовірність того, що за розглянутий період по колу може проходити електричний струм.



3.При переливанні крові слід враховувати групу крові донора і хворого. Людині, що має четверту групу крові, можна переливати кров довільної групи; людині з другою або третьою групою крові можна переливати кров або тої ж групи, або першої; людині з першою групою крові можна переливати лише кров першої групи. Серед населення 30% мають першу, 35% – другу, 25 третю і 10% – четверту групи крові. Знайти ймовірність того, що випадково взятому хворому можна перелити кров випадково взятого донора.

Глава 2. Випадкова величина

2.1.Випадкова величина

В принципі, кожна характеристика реального світу може розглядатись як випадкова величина. Її випадковість може бути пов'язана з її природою: координати і імпульси мікрочастинок. Її випадковість може бути пов'язана і з принциповою неможливістю точного її вимірювання: відстань між точками А і В. Лише в деяких випадках такі характеристики відомі точно: кількість об'єктів у фіксованому об'ємі, якщо мова йде про невелику їх кількість.

З математичної точки зору випадковою величиною називається довільна дійсна числова функція $\xi = \xi\left(\omega\right)$, визначена на просторі випадкових подій, де $\omega \in \Omega$. Нагадаємо, що ймовірність також визначена на просторі елементарних подій. Але крім того, вона визначена і на алгебрі подій, тобто на довільних підмножинах простору елементарних подій, замкнених відносно чотирьох операцій над множинами. Фактично це означає, що кожній елементарній події ми в той чи інший спосіб ставимо у відповідність дійсне число. У випадку одноразового підкидання монети простір елементарних подій складається з двох подій ω_{Γ} - випав герб і ω_{P} - випала решка $\Omega = \{\omega_{\Gamma}, \omega_{\mathrm{P}}\}$. Ми можемо, наприклад, гербу поставити у відповідність 1, а решці – 0 $1 = \xi(\omega_{\Gamma})$, $0 = \xi(\omega_{\mathrm{P}})$, сформувавши випадкову величину з наступним набором можливих значень $\xi = \{0,1\}$. Причому, це лише один з нескінченої кількості способів формування випадкової величини у даному випадку.

З максимально можливою повнотою неперервна випадкова величина описується своєю функцією розподілу

$$F_\xi(x) = P(\xi < x).$$

Ця функція розподілу має сенс ймовірності того, що випадкова величина ξ приймає значення менші від x, тобто належить інтервалу $(-\infty, x)$. Через функцію розподілу ймовірностей можна виразити і ймовірність того, що значення випадкової величини ξ належать скінченому інтервалу $[x_1, x_2]$. Дійсно, розглядаючи інтервал $(-\infty, x_2)$ або $(\xi < x_2)$ як подію, що належить простору елементарних подій $\Omega = (-\infty, \infty)$ або $\Omega = (-\infty < \xi < \infty)$, ми можемо її представити у вигляді суми двох несумісних подій $(-\infty, x_1)$, або $(\xi < x_1)$ і (x_1, x_2), або $(x_1 \le \xi < x_2)$. Тобто

$$(-\infty, x_2) = (-\infty, x_1) + (x_1, x_2),$$

або

$$(\xi < x_2) = (\xi < x_1) + (x_1 \le \xi < x_2).$$

Відповідно до аксіоми адитивності ймовірностей

$$P(\xi < x_2) = P[(\xi < x_1) + (x_1 \le \xi < x_2)] =$$
$$= P(\xi < x_1) + P(x_1 \le \xi \le x_2).$$

Мовою функцій розподілу ймовірностей останню рівність можна записати так

$$F_\xi(x_2) = F_\xi(x_1) + P(x_1 \le \xi < x_2).$$

Звідси шукана ймовірність

$$P(x_1 \le \xi < x_2) = F_\xi(x_2) - F_\xi(x_1).$$

Функція розподілу має наступні властивості:

1.Якщо $x_1 \le x_2$, то $F_\xi(x_1) \le F_\xi(x_2)$. Ця властивість означає, що функція розподілу є монотонно зростаючою функцією;

2. $\lim\limits_{x \to -\infty} F_{\xi}\left(x\right) = F_{\xi}\left(-\infty\right) = 0; \lim\limits_{x \to \infty} F_{\xi}\left(x\right) = F_{\xi}\left(\infty\right) = 1$. Ця властивість є наслідком відповідної аксіоми ймовірності, про те що ймовірність неможливої події дорівнює нулю, а ймовірність достовірної події дорівнює одиниці;

3. $\lim\limits_{x \to x_0 - 0} F_{\xi}\left(x\right) = F_{\xi}\left(x_0\right)$. Ця властивість вимагає від функції розподілу неперервності зліва в кожній точці її області визначення. При цьому неперервність функції в кожній точці не вимагається. Нагадаємо, що для неперервності функції в даній точці мають існувати і збігатись ліва і права межі функції.

Фактично функція розподілу ймовірності є інтегральною характеристикою випадкової величини, оскільки містить всю можливу інформацію про випадкову величину на проміжку $\left(-\infty, x\right)$. Становить великий інтерес локальна інформація про випадкову величину, тобто інформація про її поведінку в нескінченно малому околі точки x. Таку локальну інформацію про випадкову величину містить у собі інша характеристика випадкової величини $P_{\xi}\left(x\right)$, яка називається густиною розподілу. Вона пов'язана з функцією розподілу, або інтегральним співвідношенням

$$F_{\xi}\left(x\right) = \int_{-\infty}^{x} P_{\xi}\left(y\right) dy,$$

або диференційним співвідношенням

$$\frac{dF_{\xi}\left(x\right)}{dx} = P_{\xi}\left(x\right).$$

Остання рівність має місце там, де функція $F_{\xi}\left(x\right)$ неперервна. Густина розподілу $P_{\xi}\left(x\right)$ має простий математичний сенс

$$P_{\xi}\left(x\right) dx = dP$$

є ймовірністю того, що значення випадкової величини ξ лежать в інтервалі $\left[x, x+dx\right]$, тобто в безпосередньому околі точки x.

Густина розподілу ймовірностей має наступні властивості:

1. $P_\xi\left(x\right) \geq 0$, $x \in \left(-\infty, \infty\right)$. Це властивість невід'ємності густини ймовірностей. Враховуючи, що ймовірність – невід'ємна величина безпосередньо пов'язана з густиною ймовірності, ця властивість є цілком природною;

2. $\int\limits_{-\infty}^{\infty} P_\xi\left(x\right)dx = 1$. Це умова нормування густини розподілу ймовірностей. Вона є наслідком відповідної властивості функції розподілу

$$F_\xi\left(\infty\right) = \int\limits_{-\infty}^{\infty} P_\xi\left(x\right)dx = 1.$$

Останнє співвідношення визначає і локальні властивості функції $P_\xi\left(x\right)$. Останні мають бути такими, щоб умова нормування виконувалась. Наприклад, функція $P_\xi\left(x\right)$ може мати скінчену кількість розривів першого роду на довільному, але скінченому проміжку. Зауважимо, що розриви першого роду можуть бути і у функції $F_\xi\left(x\right)$. Через густину розподілу дуже зручно виразити ймовірність приналежності значень випадкової величини певному інтервалу. Дійсно

$$P\left(a \leq x < b\right) = \int\limits_a^b P_\xi\left(x\right)dx.$$

Останнє співвідношення називають законом розподілу неперервної випадкової величини. Таким чином закон розподілу відповідає на запитання: з якою ймовірністю значення випадкової величини лежать в тому чи іншому інтервалі.

Зазначимо, що для неперервної випадкової величини не має сенсу питання про ймовірність конкретного значення випадкової величини, оскільки ця ймовірність завжди є нескінченно малою величиною.

Для дискретної випадкової величини має сенс постановка питання про ймовірність кожного окремого її значення і закон розподілу дискретної випадкової величини завжди можна записати у вигляді таблиці з двох рядків або стовпців. Один з них - можливі значення випадкової величини

$$x_1, x_2, ..., x_N,$$

інший – ймовірності цих значень

$$P_1, P_2, ..., P_N,$$

де

$$P_i = P\left(\xi = x_i\right).$$

Отже, випадкова величина – це доволі складний математичний об'єкт, що має дві структурні складові: інтервал можливих значень і функцію розподілу ймовірностей або густину розподілу ймовірностей цих значень для неперервної випадкової величини; набір можливих значень та ймовірностей цих значень для дискретної випадкової величини.

Задачею математичної статистики є отримання максимально можливої інформації про випадкову величину шляхом її вимірювання.

Нехай задані декілька випадкових величин $\xi_1 = \xi_1\left(\omega\right)$, $\xi_2 = \xi_2\left(\omega\right), \ldots,$ $\xi_r = \xi_r\left(\omega\right)$, де $\omega \in \Omega$. У цьому разі кожній елементарній події ставиться у відповідність не одне число, а r – вимірний вектор

$$\xi = \left(\xi_1, ..., \xi_r\right).$$

Сумісною функцією розподілу ймовірностей або багатовимірною функцією розподілу неперервної випадкової величини називається наступна ймовірність

$$F_\xi\left(x\right)=F_{\xi_1\xi_2...\xi_r}\left(x_1,x_2,...x_r\right)=P\left(\xi_1<x_1,\xi_2<x_2,...,\xi_r<x_r\right),$$

що розглядається як функція точки $x=\left(x_1,...,x_r\right)$ r – вимірного евклідового простору R^r.

Якщо між окремими випадковими величинами $\xi_1,...,\xi_r$ відсутній причинно-наслідковий зв'язок, то такі випадкові величини є незалежними. Для них

$$F_{\xi_1,\xi_2,...,\xi_r}(x_1,x_2,...,x_r)=F_{\xi_1}(x_1)F_{\xi_2}(x_2)...F_{\xi_r}(x_r),$$

тобто багатовимірна функція розподілу є добутком одновимірних функцій розподілу. Багатовимірна функція містить, фактично, ту ж інформацію про багатовимірну випадкову величину, що і одновимірна про одновимірну випадкову величину. Аналогічну властивість має і багатовимірна густина розподілу

$$P_{\xi_1,...\xi_r}\left(x_1,...,x_r\right)=P_{\xi_1}\left(x_1\right)...P_{\xi_r}\left(x_r\right).$$

Для дискретного закону розподілу

$$P\left(\xi_1=x_1,...,\xi_r=x_r\right)=P\left(\xi_1=x_1\right)...P\left(\xi_r=x_r\right).$$

Крім випадкових величин можна розглядати і довільні функції випадкових величин, які, очевидно, також будуть випадковими величинами. Якщо випадкові величини $\xi_1,...,\xi_r$ є незалежними, то довільні функції цих випадкових величин $\varphi_1\left(\xi_1\right),...,\varphi_r\left(\xi_r\right)$ також будуть незалежними випадковими величинами. Можна довести і загальніше твердження, що незалежними будуть і функції

$$f_1\left(\xi_1,...,\xi_r\right),...,f_r\left(\xi_1,...,\xi_r\right),....$$

Закон розподілу нових випадкових величин є певною функцією закону розподілу старих випадкових величин.

Розглянемо спочатку випадок одновимірної випадкової величини ξ і функції цієї випадкової величини $\eta=\varphi\left(\xi\right)$. Нехай функція розподілу випадкової величини ξ є відомою. Знайдемо функцію розподілу випадкової величини η. Очевидно, що

$$F_\eta(x) = P(\eta < x) = P\left[\varphi(\xi) < x\right].$$

Умову $\varphi(\xi) < x$ можна записати як $\xi < \varphi^{-1}(x)$, де $\varphi^{-1}(x)$ функція обернена щодо функції $\varphi(x)$. Тоді

$$F_\eta(x) = P\left[\eta < \varphi^{-1}(x)\right].$$

Ймовірність у правій частині останньої рівності можна обчислити, якщо відома функція розподілу $F_\xi(x)$.

Розглянемо тепер багатовимірний випадок. Нехай $\eta = \eta(\eta_1, \eta_2, ..., \eta_r)$ і $\eta = g(\xi)$ де $\xi = (\xi_1, ..., \xi_r)$. Тут ξ і η – випадкові вектори. Нехай при цьому відома багатовимірна густина розподілу $P_\xi(x)$, де $x = (x_1, ..., x_r)$, $g(x) = (g_1(x), ..., g_r(x))$ – вектор-функція. Якщо відображення $y = g(x)$ взаємно однозначне, неперервне диференційоване і якобіан

$$I(x) = \frac{D(g_1, ..., g_r)}{D(x_1, ..., x_r)} \neq 0,$$

то багатовимірна густина розподілу вектора η визначається так

$$P_\eta(x) = P_\xi\left[g^{-1}(x)\right] / I\left[g^{-1}(x)\right],$$

де $g^{-1}(x)$ – вектор-функція, обернена вектор-функції $g(x)$.

Детальніше функції розподілу ймовірностей багатьох випадкових величин ми розглянемо у четвертій главі.

Задача 1. *Нехай випадкова величина ξ має нормальний розподіл з параметрами (a, σ^2). Знайдемо закон розподілу випадкової величини $\eta = A\xi + B$.*

Розв`язання. *Нехай спочатку* $A>0$. *Тоді*

$$F_\eta(x)=P(A\xi+B<x)=P\left(\xi<\frac{x-B}{A}\right)=\int_{-\infty}^{(x-B)/A}P_\xi(y)dy.$$

Отже функцію розподілу випадкової величини η *ми знайшли. Знайдемо тепер її густину розподілу. За означенням густини розподілу випадкової величини* η

$$P_\eta(x)=\frac{dF_\eta(x)}{dx}=P_\xi\left(\frac{x-B}{A}\right)\left(\frac{x-B}{A}\right)'=\frac{1}{A}P_\xi\left(\frac{x-B}{A}\right).$$

Нехай тепер $A<0$. *Тоді функція розподілу буде дещо іншою*

$$F_\eta(x)=P(A\xi+B<x)=P\left(\xi>\frac{x-B}{A}\right)=\int_{(x-B)/A}^{\infty}P_\xi(y)dy.$$

Для густини розподілу аналогічно попередньому матимемо

$$P_\eta(x)=\frac{dF_\eta(x)}{dx}=-P_\xi\left(\frac{x-B}{A}\right)\left(\frac{x-B}{A}\right)'=-\frac{1}{A}P_\xi\left(\frac{x-B}{A}\right)=\frac{1}{|A|}P_\xi\left(\frac{x-B}{A}\right).$$

Оскільки нам відомо, що

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

то як для $A>0$, *так і для* $A<0$

$$P_\eta(x)=\frac{1}{|A|\sigma\sqrt{2\pi}}\exp\left[-\frac{\left(\frac{x-B}{A}-a\right)^2}{2\sigma^2}\right]=\frac{1}{\tilde{\sigma}\sqrt{2\pi}}\exp\left[-\frac{(x-\tilde{a})^2}{2\tilde{\sigma}^2}\right],$$

де $\tilde{\sigma}=|A|\sigma$, $\tilde{a}=Aa+B$.

.

Задача 2. *Нехай* $h(x)=\cos(x)$, *а густина розподілу ймовірності*

$$f_{\xi}(x)=\begin{cases} a+bx,\ x\in[0,\pi/2], \\ \quad 0,\ x\notin[0,\pi/2], \end{cases}$$

де $a=1/\pi$, $b=4/\pi^2$. *Знайти густину розподілу ймовірності для випадкової величини* h.

Розв'язання. *Очевидно головне значення оберненої функції* $x=\omega\cos(h)$ *є функцією однозначною. Отже,*

$$g_{\eta}(h)=\left[a+b\omega\cos(h)\right]\left|\frac{da\cos(h)}{dh}\right|=$$

$$=\left[a+b\omega\cos(h)\right]\frac{1}{\sqrt{1-h^2}}.$$

Задача 3. *Нехай* $h(x)=x^2$, *а густина розподілу ймовірності*

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

Знайти густину розподілу ймовірності для випадкової величини h.

Розв`язання. *Очевидно обернена функція щодо функції* $h(x)=x^2$ *є двозначною. Її можна представити сукупністю наступних двох гілок*

$$x_1(h)=\sqrt{h},\ x_2(h)=-\sqrt{h}.$$

Відповідно,

$$\left|\frac{dx_1(h)}{dh}\right|=\left|\frac{dx_2(h)}{dh}\right|=\frac{1}{2\sqrt{h}}.$$

Тепер

$$g_\eta(x)=\frac{1}{\sqrt{2\pi h}}\exp(-h/2).$$

Останній результат є частинним випадком гама - розподілу

$$g_\eta(h)=\frac{1}{\alpha!\beta^{\alpha+1}}h^\alpha e^{-h/\beta},\ 0\le h<\infty,$$

де $\alpha=-1/2$, $\beta=2$.

2.2.Математичне очікування

Максимально повна інформація про випадкову величину міститься в її функції розподілу. Однак, на практиці часто цілком достатньою і меншої інформації про випадкову величину. Бажано, щоб така інформація мала вигляд дійсних чисел – єдиних математичних об’єктів, які можна безпосередньо порівнювати між собою. Однією з таких числових характеристик є математичне очікування. Математичне очікування – це один з варіантів середнього значення випадкової величини, яке обчислюється із урахуванням ймовірності кожного її значення.

Для дискретної випадкової величини ξ, можливими значеннями якої є $x_1,x_2,...,x_n$, що реалізуються з ймовірностями $P_1,P_2,...,P_n$, математичне очікування визначається так

$$M\xi=\sum_{k=1}^{n}x_kP_k\ .$$

З виразу видно, що у цій сумі вагою кожного значення випадкової величини є його ймовірність. Іншим варіантом середнього значення випадкової величини є середнє арифметичне

$$\overline{\xi}=\frac{1}{n}\sum_{k=1}^{n}x_k\ .$$

У цьому виразі ігнорується той факт, що деякі значення випадкової величини при вимірюванні можуть зустрічатись рідко, а деякі – часто.

Ці два варіанти середнього значення збігаються в двох випадках. У першому з них, що реалізується для скінченого n збіг можливий лише тоді, коли всі значення випадкової величини ξ однаково ймовірні, тобто

$$P_1 = P_2 = ... = P_n = \frac{1}{n}.$$

У цьому разі

$$\mathrm{M}\xi = \overline{\xi}.$$

Другий випадок реалізується для довільних ймовірностей $P_1, P_2, ..., P_n$, але лише для $n \to \infty$. Останнє твердження є наслідком закону великих чисел.

Для неперервної випадкової величини математичне очікування визначається через густину розподілу так

$$M\xi = \int_{-\infty}^{\infty} x P_{\xi}(x)\, dx.$$

Для n – вимірної випадкової величини $\xi = (\xi_1, \xi_2, ..., \xi_n)$

$$M\xi = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \xi(x_1, x_2, ..., x_n) P_{\xi_1 ... \xi_n}(x_1, x_2, ..., x_n)\, dx_1 dx_2 ... dx_n.$$

Очевидно, що математичне очікування існує лише у разі, коли збігаються відповідна сума або інтеграли.

Задача 1. *Двоє гравців підкидають монету. Перший з них виграє 1 грн., якщо випадає герб, і програє 1 грн., якщо випадає решка. Знайти математичне очікування виграшу цього гравця.*

Розв`язання. *Простір елементарних подій складається з двох елементарних подій* ω_Γ *і* ω_P*:* $\Omega = \{\omega_\Gamma, \omega_P\}$. *Розглянемо* *наступну випадкову величину $\xi(\omega)$, яка приймає наступні значення*

$$x_1 = \xi(\omega_{\Gamma}) = 1,$$

$$x_2 = \xi(\omega_P) = -1.$$

Очевидно, ця випадкова величина описує можливий виграш або програш першого гравця. Для симетричної монети вірне класичне означення ймовірності з наступними елементарними ймовірностями

$$P_1 = P(\omega_{\Gamma}) = \frac{1}{2},$$

$$P_2 = P(\omega_P) = \frac{1}{2}.$$

Шукане математичне очікування буде таким

$$M\xi = x_1P_1 + x_2P_2 = 1\cdot\frac{1}{2} + (-1)\cdot\frac{1}{2} = 0.$$

Зрозуміло, що при одноразовому підкиданні монети перший гравець або виграє 1 грн., або програє 1. грн. Знайдене математичне очікування означає, що при достатньо довгій грі (нескінченно великій кількості підкидань монети) середній виграш першого гравця буде нульовим. Таким же він буде і для другого гравця.

Математичне очікування функції випадкової величини визначається подібно до математичного очікування самої випадкової величини. Для довільної функції дискретної випадкової величини

$$M\eta = \sum_{k=1}^{n} g(x_k)P_k,$$

для довільної функції неперервної випадкової величини

$$M\eta = \int_{-\infty}^{\infty} g(x)P_{\xi}(x)\,dx,$$

де

$$\eta = g\left(\xi\right).$$

Задача 2. *Обчислимо математичне очікування неперервної випадкової величини з нормальним законом розподілу*

$$P_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp-\left[\frac{\left(x-a\right)^2}{2\sigma^2}\right].$$

Розв`язання. *За означенням*

$$M\xi = \int\limits_{-\infty}^{\infty} xP_\xi\left(x\right)dx = \frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\exp\left[-\frac{\left(x-a\right)^2}{2\sigma^2}\right]xdx\,.$$

Виконаємо підстановку $x-a=y$*, тоді*

$$M\xi = \frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\exp\left(-\frac{y^2}{2\sigma^2}\right)ydy + \frac{a}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\exp\left(-\frac{y^2}{2\sigma^2}\right)dy\,.$$

Перший інтеграл у правій частині дорівнює нулю, оскільки його підінтегральна функція є непарною, а границі інтегрування – симетричними. Другий інтеграл є інтегралом Пуассона

$$\int\limits_{-\infty}^{\infty}\exp\left(-\alpha x^2\right)dx = \sqrt{\frac{\pi}{\alpha}}\,.$$

Отже математичне очікування є одним з параметрів нормального розподілу

$$\mathrm{M}\xi = a\,.$$

Задача 3. *Обчислимо математичне очікування неперервної випадкової величини з показниковим законом розподілу*

$$P_\xi\left(x\right) = \alpha\exp\left(-\alpha x\right),\ x \ge 0\,.$$

Розв`язання. *За означенням*

$$M\xi = \int\limits_{0}^{\infty} xP_\xi(x)dx = \alpha\int\limits_{0}^{\infty} x\exp(-\alpha x)dx\,.$$

Інтегруючи частинами маємо

$$M\xi = -x\exp\left(-\alpha x\right)\Big|_0^\infty + \int\limits_0^\infty \exp\left(-\alpha x\right)dx = \frac{1}{\alpha}.$$

Задача 4. *Обчислимо математичне очікування дискретної випадкової величини з біноміальним розподілом*

$$P_n\left(m\right) = C_n^m p^m \left(1-p\right)^{n-m}.$$

Розв`язання. *За означенням*

$$M\xi = \sum_{m=0}^{n} mP_n\left(m\right) = \sum_{m=0}^{n} mC_n^m p^m \left(1-p\right)^{n-m} =$$

$$= \sum_{m=0}^{n} m\frac{n\left(n-1\right)...\left(n-m+1\right)}{m!} p^m \left(1-p\right)^{n-m} =$$

$$= pn\sum_{m=1}^{n} \frac{\left(n-1\right)...\left(n-m+1\right)}{\left(m-1\right)!} p^{m-1} \left(1-p\right)^{n-1-\left(m-1\right)} =$$

$$= pn\sum_{k=0}^{n-1} \frac{\left(n-1\right)...\left(n-1-k+1\right)}{k!} p^k \left(1-p\right)^{n-1-k} = p\cdot n.$$

Оскільки умова нормування

$$\sum_{k=0}^{n-1} C_{n-1}^k p^k \left(1-p\right)^{n-1-k} = 1.$$

Задача 5. *Обчислимо математичне очікування дискретної випадкової величини з Пуасонівським розподілом.*

Розв`язання. *За означенням*

$$M\xi = \sum_{m=0}^{\infty} m\frac{\lambda^m}{m!}\exp\left(-\lambda\right) = \lambda\exp\left(-\lambda\right)\sum_{m=0}^{\infty}\frac{\lambda^{m-1}}{\left(m-1\right)!} =$$

$$= \lambda\exp\left(-\lambda\right)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = \lambda.$$

Отже математичне очікування є параметром розподілу Пуассона.

Математичне очікування має наступні властивості:

1.Якщо C – довільна стала, то $MC = C$.

2.Якщо C – довільна стала, то $\mathrm{M} C\xi = C\mathrm{M}\xi$.

3.Для довільної випадкової величини ξ

$$\left|M\xi\right| \le M\left|\xi\right|.$$

Очевидно знак рівності буде лише у разі, коли випадкова величина ξ приймає лише додатні значення.

4.Для довільних випадкових величин ξ_1 і ξ_2

$$M\left(\xi_1 + \xi_2\right) = M\xi_1 + M\xi_2.$$

5.Для незалежних випадкових величин ξ_1 і ξ_2

$$M(\xi_1 \xi_2) = M(\xi_1)M(\xi_2).$$

Зауваження: Якщо випадкові величини ξ_1 *і* ξ_2 *довільні, зокрема залежні, то останнє співвідношення матиме вигляд*

$$M\xi_1\xi_2 = M\xi_1 M\xi_2 + \mathrm{cov}\left(\xi_1, \xi_2\right),$$

де останній доданок називається коваріацією (див. далі).

6.Якщо при всіх значеннях випадкових величин ξ і η $\xi \ge \eta$, то $M\xi \ge M\eta$.

Всі наведені властивості є наслідком відповідних властивостей сум та інтегралів. Наступна властивість математичного очікування є не такою очевидною

7. $P(\xi \ge \varepsilon) \le M\xi / \varepsilon$, де випадкова величина $\xi \ge 0$ і має скінчене математичне очікування ($\varepsilon > 0$).

Доведення. *Практичний сенс остання нерівність має у разі, коли* $\varepsilon > M\xi$ *і є наслідком попередньої нерівності при спеціальному виборі випадкової величини* η. *Отже, нехай областю визначення випадкової величини* ξ *буде весь простір елементарних подій* Ω. *Виділимо з нього підмножину* Ω_ε, *де функція* $\xi(\omega) \ge \varepsilon$. *Причому, на іншій підмножині* $\Omega - \Omega_\varepsilon$ $\xi(\omega) < \varepsilon$

. *Сформуємо тепер іншу випадкову величину* $\eta = 0$ *на підмножині* $\Omega - \Omega_\varepsilon$ *і* $\eta = \varepsilon$ *на підмножині* Ω_ε. *У цієї випадкової величини дві властивості. Перша – для неї має місце нерівність* $\xi \ge \eta$, *а отже і властивість* $M\xi \ge M\eta$. *Друга – математичне очікування цієї випадкової величини легко обчислюється*

$$M\eta = \int_{\Omega} x(\omega)P_\eta[x(\omega)]dx(\omega) =$$

$$= \int_{\Omega_\varepsilon} xP_\eta(x)dx = \varepsilon \int_{\Omega_\varepsilon} P_\eta(x)dx = \varepsilon P(\xi \ge \eta).$$

Отже,

$$M\xi \ge \varepsilon P(\xi \ge \varepsilon)$$

або

$$P(\xi \ge \varepsilon) \le \frac{M\xi}{\varepsilon}.$$

Зауважимо, що $M\xi = 0$ *за умови, що* $\xi \ge 0$, *лише у разі,* $\xi = 0$.

Задача 6. *Нехай випадкова величина* ξ *має нормальний розподіл з параметрами* $\left(a, \sigma^2\right)$. *Знайдемо математичне очікування величини* $\eta = A\xi + B$.

Розв`язання. *Очевидно*

$$\mathrm{M}\eta = \mathrm{M}\left(A\xi + B\right) = A\mathrm{M}\xi + B = Aa + B,$$

де використали відомий за умовою факт, що $\mathrm{M}\xi = a$.

Задача 7. *Нехай неперервна випадкова величина має рівномірний розподіл*

$$P_\xi\left(x\right) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & x < 0 \cup x > b \end{cases}.$$

Знайдемо її математичне очікування.

Розв`язання. *За означенням*

$$M\xi = \frac{1}{b-a}\int_a^b x dx = \frac{a+b}{2}.$$

2.3.Формули Шенона і Хартлі

Інформація про певну подію, з точки зору сучасної математики, напряму пов'язана з ймовірністю реалізації цієї події. Відповідно американському математику Клоду Шенону кількість інформації в бітах про подію A визначається так

$$I(A) = -\log_2[P(A)].$$

Тут $P(A)$ - ймовірність настання події A. Оскільки ймовірність не перевищує одиницю, то логарифм такого аргументу завжди буде від'ємним. Для того, щоб кількість інформації завжди визначалась додатним числом, потрібно перед логарифмом поставити знак мінус. Основу 2 для логарифму Клод Шенон обрав для того, щоб кількість інформації в найпростішому повідомленні дорівнювала 1, точніше 1 біту. Дійсно, найпростіше повідомлення стосується випадкового експерименту з двома рівноправними наслідками. Кожний наслідок тоді реалізується з ймовірністю 1/2. Відповідно, інформація про настання будь-якого з цих наслідків $I = -\log_2(1/2) = 1$ біт.

Задача 1. *Дві симетричні монети підкидаються один раз. Скільки інформації містить повідомлення, що на обох монетах випав герб?*

Розв'язання. Ймовірність випадення двох гербів $P_{\text{ГГ}} = 1/4/$

Відповідно $I = -\log_2(1/4) = 2$ біт.

Логарифм Клод Шенон використав для того, щоб інформація про незалежні повідомлення була адитивною, тобто щоб кількість інформації в повідомленні про настання двох незалежних подій A і B дорівнювала сумі кількостей інформації, що окремо міститься в кожному повідомленні

$$I(A \cdot B) = I(A) + I(B).$$

Логарифм якраз і має відповідну властивість

$$\log_2[P(A \cdot B)] = \log_2[P(A) \cdot P(B)] = \log_2[P(A)] + \log_2[P(B)].$$

Кожний випадковий експеримент має декілька наслідків. Повідомлення можуть стосуватись настання кожного з них. Можна говорити про середню кількість інформації, що міститься в повідомленні для серії повідомлень, що охоплюють всі можливі наслідки даного випадкового експерименту. Очевидно, що це є математичне очікування кількості інформації, що враховує кількість інформації в кожному повідомленні, можливого щодо даного випадкового експерименту. Це і є знаменита формула Клода Шенона

$$I = -\sum_{k=1}^{n} P(A_k) \log_2[P(A_k)].$$

З'ясуємо, скільки інформації міститься у відповіді на питання з двома можливими наслідками (так або ні). Нехай P – ймовірність того, що на поставлене запитання буде відповідь «так». Тоді ймовірність відповіді «ні» є $q = 1 - P$. Відповідно до формули Шенона, середня кількість інформації, що міститься в двох можливих повідомленнях

$$I = -P \log_2 P - (1 - P) \log_2 (1 - P).$$

Якщо $P = 0$, то $q = 1$ і $I = 0$. Тобто інформація, отримана з відповіді, дорівнює нулю. Такою ж вона буде і у протилежному випадку $P = 1$, то $q = 0$. Максимальною і рівною одиниці (одному біту) вона буде у разі рівноправності обох відповідей, максимальній невизначеності ситуації, тобто у разі, коли $P = q = 1/2$ $(I = \log_2 2 = 1)$.

Інформація, що міститься в одному повідомленні, середня кількість інформації, щ міститься в усій сукупності можливих повідомлень, збігаються лише в одному випадку, коли всі наслідки випадкового експерименту є рівноправними, тобто ймовірності цих наслідків є однаковими. Тоді

$$I = -\sum_{k=1}^{n} P(A_k)\log_2[P(A_k)] = \log_2(n),$$

оскільки всі $P(A_k) = 1/n$. При цьому і

$$I(A_k) = -\log_2[P(A_k)] = \log_2(n).$$

Отже,

$$I(A_k) = I.$$

Такий спрощений варіант формули Шенона є не менш знаменитим і має назву формула Хартлі.

Важливим застосуванням формул Шенона і Хартлі є їх використання для аналізу результатів вимірювань. Нехай ζ – дискретна випадкова величина, що може приймати значення $x_1, x_2, ..., x_N$ із ймовірностями $P_1, P_2, ..., P_N$. Ми вимірюємо цю випадкову величину, отримуючи одне з можливих її значень. Яку інформацію ми при цьому отримуємо? Відповідно американському математику Клоду Шенону кількість інформації в бітах визначається наступним математичним очікуванням

$$I(\zeta) = -\sum_{n=1}^{N} P_n \log_2 P_n .$$

Якщо всі наслідки рівноправні, то $P_n = 1/N$. Тоді

$$I(\zeta) = \log_2 N$$

і ми одержуємо формулу Хартлі

2.4.Дисперсія

Математичне очікування випадкової величини містить досить суттєву, але не повну інформацію про випадкову величину. Відповідаючи на питання якою є середнє значення випадкової величини, вона нічого не говорить про розкиданість значень випадкової величини відносно цього середнього значення. На останнє запитання відповідає інша характеристика, яка називається дисперсією. За означенням

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

Тобто дисперсія, характеризуючи розкиданість значень випадкової величини відносно її середнього значення, є математичним очікуванням квадрату відхилення значень випадкової величини відносно її середнього значення. Цю розкиданість можна схарактеризувати у безліч способів але запропонований є найпростішим. Виразу для дисперсії можна надати і іншого вигляду

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

Дійсно,

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

$$= M\xi^2 - 2M\left(\xi M\xi\right) + M\left(M\xi\right)^2 =$$

$$= M\xi^2 - 2M\xi \cdot M\xi + \left(M\xi\right)^2 = M\xi^2 - \left(M\xi\right)^2 .$$

Дисперсія має наступні властивості:

1.Для довільної випадкової величини ξ дисперсія є невід’ємною величиною $D\xi \geq 0$.

2.Якщо C – довільна стала, то $DC = 0$.

3.Якщо C – довільна стала, то $DC\xi = C^2 D\xi$.

4.Для незалежних випадкових величин ξ_1 і ξ_2

$$D\left(\xi_1+\xi_2\right)=D\xi_1+D\xi_2 .$$

5. $P(|\,\xi - M\xi\,| \geq \varepsilon) \leq \dfrac{D\xi}{\varepsilon^2}$, якщо випадкова величина ξ має скінчену дисперсію.

Доведення. *Якщо виконується нерівність* $|\,\xi - M\xi\,| \geq \varepsilon$, *то виконується і нерівність* $(\xi - M\xi)^2 \geq \varepsilon^2$. *Взявши в якості невід'ємної випадкової величини* $\eta = |\,\xi - M\xi\,|$ *і врахувавши, що* $\eta^2 = (\xi - M\xi)^2$, *а* $M\eta^2 = M(\xi - M\xi)^2 = D\xi$ *на основі відповідної властивості математичного очікування отримаємо*

$$P(|\,\xi - M\xi\,| \geq \varepsilon) = P(\eta \geq \varepsilon) = .$$

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

Остання нерівність має практичну користь лише у разі, якщо $D\xi < \varepsilon^2$. *При цьому випадкова величина* ξ *може набувати і додатних, і від'ємних значень.*

Обчислимо дисперсію деяких випадкових величин.

Задача 1. *Нехай неперервна випадкова величина має нормальний розподіл. Знайдемо її дисперсію.*

Розв`язання. *За означенням*

$$D\xi = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(x - M\xi\right)^2 \exp\left[-\frac{\left(x-a\right)^2}{2\sigma^2}\right] dx =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(x - a\right)^2 \exp\left[-\frac{\left(x-a\right)^2}{2\sigma^2}\right] dx .$$

Для обчислення використаємо інтеграл Пуассона

$$\int_{-\infty}^{\infty} \exp\left(-\alpha x^2\right) dx = \sqrt{\frac{\pi}{\alpha}} .$$

Диференціюємо його за параметром α

$$\int_{-\infty}^{\infty} x^2 \exp\left(-\alpha x^2\right) dx = \frac{\pi^{1/2}}{2\alpha^{3/2}}.$$

У висліді

$$D\xi = \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{\sqrt{\pi}\, 2^{3/2} \sigma^3}{2} = \sigma^2.$$

Отже дисперсія також є параметром нормального розподілу.

Задача 2. *Нехай неперервна випадкова величина розподілена рівномірно. Знайдемо її дисперсію.*

Розв`язання. *За означенням*

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

Оскільки, як було показано вище,

$$M\xi = \frac{a+b}{2},$$

то

$$D\xi = \int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 P_\xi\left(x\right) = \frac{1}{b-a} \int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 dx =$$

$$= \frac{1}{b-a} \cdot \frac{1}{3} \left(x - \frac{a+b}{2}\right) \Bigg|_a^b = \frac{\left(b-a\right)^2}{12}.$$

Задача 3. *Нехай дискретна випадкова величина розподілена за законом Пуассона. Знайдемо її дисперсію.*

Розв`язання. *За означенням*

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

Останній вираз зручно записати так

$$D\xi = M\xi\left(\xi - 1\right) + M\xi - \left(M\xi\right)^2.$$

Як було показано вище $M\xi = \lambda$*, тоді*

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

Обчислимо перший доданок у правій частині

$$M\xi\left(\xi-1\right)=\sum_{m=0}^{\infty}m\left(m-1\right)\frac{\lambda^m}{m!}\exp\left(-\lambda\right)=$$

$$=\lambda^2\sum_{m=2}^{\infty}\frac{\lambda^{m-2}}{\left(m-2\right)!}\exp\left(-\lambda\right)=\lambda^2\exp\left(-\lambda\right)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=\lambda^2,$$

оскільки

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=\exp\left(\lambda\right).$$

Остаточний результат буде таким

$$D\xi=\lambda.$$

Отже, дисперсія дискретної випадкової величини, розподіленої за законом Пуассона, є параметром цього розподілу. Зауважимо, що характерною рисою цього закону є збіг математичного очікування і дисперсії

$$M\xi=D\xi=\lambda.$$

Задача 4. *Нехай дискретна випадкова величина розподілена за біноміальним законом. Знайдемо її дисперсію.*

Розв`язання. *За означенням*

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

Останній вираз зручно записати так

$$D\xi=M\xi\left(\xi-1\right)+M\xi-\left(M\xi\right)^2.$$

Як було показано вище $M\xi=nP$*, тоді*

$$D\xi=M\xi\left(\xi-1\right)+np-n^2p^2.$$

Обчислимо перший доданок у правій частині

$$M\xi\left(\xi-1\right)=\sum_{m=0}^{n}m\left(m-1\right)\frac{n\left(n-1\right)...\left(n-m+1\right)}{m!}p^m\left(1-p\right)^{n-m}=$$

$$= n(n-1)\, p^2 \sum_{m=2}^{n} \frac{(n-2)\ldots\left[n-2-(m-2)+1\right]}{(m-2)!} p^{m-2} (1-p)^{n-2-(m-2)} =$$

$$= n(n-1)\, p^2 \sum_{k=0}^{n-2} \frac{(n-2)\ldots\left[n-2-k+1\right]}{k!} p^{k} (1-p)^{n-2-k} .$$

Оскільки, відповідно до умови нормування

$$\sum_{k=0}^{n-2} \frac{(n-2)\ldots\left[n-2-k+1\right]}{k!} p^{k} (1-p)^{n-2-k} = 1,$$

то

$$M\xi(\xi-1) = n(n-1)\, p^2 .$$

Остаточний результат буде таким

$$D\xi = np(1-p).$$

2.5.Коваріація

Максимально повну інформацію про дві випадкові величини ξ_1 і ξ_2 містить їх парна функція розподілу $F_{\xi_1\xi_2}(x_1, x_2)$. Суттєву частину цієї інформації несе числова характеристика, що називається коваріацією двох випадкових величин. Ця характеристика є кількісною характеристикою залежності або незалежності випадкових величин ξ_1 і ξ_2. За означенням

$$\operatorname{cov}(\xi_1, \xi_2) = M\,(\xi_1 - M\xi_1)(\xi_2 - M\xi_2).$$

Використовуючи властивості математичного очікування легко отримати наступну формулу

$$\operatorname{cov}(\xi_1, \xi_2) = M\xi_1\xi_2 - M\xi_1 M\xi_2 .$$

Очевидно, що якщо випадкові величини незалежні, тобто $M\xi_1\xi_2 = M\xi_1 M\xi_2$, то $\operatorname{cov}(\xi_1, \xi_2) = 0$. Якщо ж вони максимально залежні, що можливо лише у разі $\xi_1 = \xi_2$, то коваріація приймає максимально можливе значення $\operatorname{cov}(\xi_1, \xi_2) = D\xi$. Серед властивостей дисперсії суми двох випадкових величин ми серед інших наводили таку, вірну лише для незалежних випадкових величин,

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2 .$$

Очевидно для довільних випадкових величин

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2 + 2\operatorname{cov}(\xi_1 + \xi_2),$$

що перевіряється безпосередньо. Так само легко перевіряються і наступні властивості:

$$\operatorname{cov}(\xi_1, \xi_2) = \operatorname{cov}(\xi_2, \xi_1),$$

$$\operatorname{cov}(C\xi_1, \xi_2) = C\operatorname{cov}(\xi_1, \xi_2),$$

де C – довільна стала.

Якщо випадкових величин довільне число n, то відповідна властивість дисперсії легко узагальнюється на цей випадок, а саме

$$D(C_1\xi_1 + C_2\,\xi_2 + \ldots + C_n\,\xi_n) = \sum_{i,j=1}^{n} \sigma_{ij} C_i C_j ,$$

де $C_1, \ldots, C_n$ – довільні сталі, а

$$\sigma_{ij} = \operatorname{cov}(\xi_i, \xi_j).$$

Можна довести, що існує наступна оцінка коваріації

$$\left|\operatorname{cov}(\xi_1, \xi_2)\right| \le \sqrt{D\xi_1 D\xi_2} ,$$

яка дозволяє ввести нову числову характеристику залежності випадкових величин – коефіцієнт кореляцій $\rho(\xi_1, \xi_2)$, а саме:

$$\rho(\xi_1, \xi_2) = \frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{D\xi_1 D\xi_2}} .$$

Зручність коефіцієнта кореляції полягає у тому, що його значення за модулем не перевищують одиницю

$$\left|\rho\left(\xi_1,\xi_2\right)\right|\le 1 .$$

При цьому, як і коваріація, коефіцієнт кореляцій двох незалежних випадкових величин $\rho\left(\xi_1,\xi_2\right)=0$. Якщо ж незалежні величини пов'язані лінійним чином

$$\xi_2 = A\xi_1 + B ,$$

де A і B – довільні сталі, то коефіцієнт кореляції приймає максимально можливе значення $\left|\rho\left(\xi_1,\xi_2\right)\right|=1$.

Зауважимо, що рівність нулю коефіцієнта кореляції є лише необхідною умовою незалежності випадкових величин.

2.6.Моменти випадкової величини

Початковим моментом k-го порядку випадкової величини ξ називається число $M\xi^k$. Центральним моментом k-го порядку випадкової величини називається число $M\left(\xi - M\xi\right)^k$. Очевидно, що довільний центральний момент k-го порядку можна представити відповідною комбінацією початкових моментів і навпаки. Отже математичне очікування випадкової величини є її початковим моментом першого порядку, а дисперсія – центральним моментом другого порядку. Кожний момент випадкової величини містить лише частину інформації про цю випадкову величину. Повну інформацію про випадкову величину містить вся нескінченна кількість або початкових, або центральних моментів. З цього правила є багато виключень. Наприклад, нормальний розподіл ймовірностей повністю визначається двома моментами – математичним очікуванням і дисперсією, що входять в якості параметрів у його густину розподілу. Тому, не випадково, на практиці найважливішими характеристиками випадкової величини є саме її математичне очікування та дисперсія.

Якщо ми маємо декілька випадкових величин $\xi_1, \xi_2, ..., \xi_n$, то крім початкових і центральних моментів, що містять інформацію про кожну з них, потрібно розглядати і всю сукупність змішаних початкових і центральних моментів. Саме змішані моменти визначають ступінь залежності різних випадкових величин. Початкові мішані моменти мають вигляд

$$M\xi_1^{k_1}\xi_2^{k_2}...\xi_n^{k_n}.$$

Центральні мішані моменти визначаються так

$$M\left(\xi_1 - M\xi_1\right)^{k_1}\left(\xi_2 - M\xi_2\right)^{k_2}...\left(\xi_n - M\xi_n\right)^{k_n}.$$

Для обчислення мішаних моментів потрібні функції розподілу відповідних сукупностей випадкових величин.

Стосовно всіх можливих моментів зазначимо, що з існування моментів k-го порядку завжди випливає існування моментів менших порядків: $k-1, k-2, ...$.

Задача 1. *Довести, що для випадкових величин ξ_1 і ξ_2, пов'язаних між собою співвідношенням $\xi_2 = A\xi_1 + B$, де A і B – довільні сталі, $\left|\rho\left(\xi_1, \xi_2\right)\right| = 1$.*

Розв`язання. *Нехай $M\xi_1 = a$, а $D\xi_1 = \sigma^2$. Тоді*

$$M\xi_2 = M\left(A\xi_1 + B\right) = M\left(A\xi_1\right) + MB = AM\xi_1 + B = Aa + B,$$

де ми використали властивості математичного очікування. Оскільки довільна стала і випадкова величина є незалежними величинами, то

$$D\xi_2 = D\left(A\xi_1 + B\right) = D\left(A\xi_1\right) + DB = A^2 D\xi_1 = A^2 + \sigma^2,$$

де ми використали властивості дисперсії, то

$$\operatorname{cov}\left(\xi_1, \xi_2\right) = M\left[\left(\xi_1 - M\xi_1\right)\left(\xi_2 - M\xi_2\right)\right] = M\left[\left(\xi_1 - a\right)\left(\xi_2 - Aa - B\right)\right] =$$

$$= M\left[\left(\xi_1 - a\right)\left(A\xi_1 + B - Aa - B\right)\right] = M\left[\left(\xi_1 - a\right)A\left(\xi_1 - a\right)\right] =$$

$$= AM\left(\xi_1 - a\right)^2 = AD\xi_1 = A\sigma^2 .$$

Отже,

$$\rho\left(\xi_1, \xi_2\right) = \frac{\operatorname{cov}\left(\xi_1, \xi_2\right)}{\sqrt{D\xi_1 D\xi_2}} = \frac{A\sigma^2}{\sqrt{A^2\sigma^2\sigma^2}} = \frac{A}{|A|}.$$

Таким чином,

$$\left|\rho\left(\xi_1, \xi_2\right)\right| = \left|\frac{A}{|A|}\right| = \frac{|A|}{|A|} = 1.$$

Задача 2. *Довести, що для двох незалежних випадкових величин* ξ_1 *і* ξ_2 $\rho\left(\xi_1, \xi_2\right) = 0$.

Розв`язання. *За означенням*

$$\operatorname{cov}\left(\xi_1, \xi_2\right) = M\xi_1\xi_2 - M\xi_1 M\xi_2 = M\xi_1 M\xi_2 - M\xi_1 M\xi_2 = 0 .$$

2.7.Твірна функція моментів

Набір початкових або центральних теоретичних моментів випадкової величини – це один з способів задати максимально можливу інформацію про випадкову величину у вигляді зліченої послідовності чисел. Це еквівалентно заданню теоретичного закону розподілу випадкової величини. Зв'язок між моментами і густиною розподілу випадкової величини можна представити в інтегральній формі. Існує і інший спосіб представити весь набір моментів однією функцією. Ця функція також містить максимально можливу інформацію про випадкову величину і називається твірною функцією моментів. Зв'язок між твірною функцією моментів і самими моментами тепер можна представити вже у диференціальній формі. Отже, функція розподілу ймовірностей або густина розподілу ймовірностей випадкової величини, твірна функція моментів випадкової величини і моменти випадкової величини – це три еквівалентні способи концентрації максимально можливої інформації про випадкову величину.

За означенням твірна функція початкових моментів випадкової величини має вигляд

$$\varphi_\xi(x) = \sum_{i=0}^{\infty} \exp(y_i x) P_i$$

для дискретної випадкової величини,

$$\varphi_\xi(x) = \int_{-\infty}^{\infty} \exp(xy) P_\xi(y) dy$$

для неперервної випадкової величини. В подальшому для економії викладу детально зупинимось лише на неперервних випадкових величинах.

Якщо експоненту розвинути в ряд Маклорена

$$\exp(xy) = 1 + xy + \frac{(xy)^2}{2!} + \ldots + \frac{(xy)^n}{n!} + \ldots,$$

то для твірної функції моментів отримаємо аналогічне розвинення

$$\varphi_\xi(x) = 1 + \alpha_1 x + \alpha_2 \frac{x^2}{2!} + \ldots + \alpha_n \frac{x^n}{n!} + \ldots.$$

Порівнюючи останній ряд з рядом Маклорена для функції $\varphi_\xi(x)$

$$\varphi_\xi(x) = 1 + \frac{d\varphi_\xi(0)}{dx} x + \frac{d^2\varphi_\xi(0)}{dx^2} \frac{x^2}{2!} + \ldots + \frac{d^n\varphi_\xi(0)}{dx^n} \frac{x^n}{n!} + \ldots,$$

Отримаємо диференційний зв’язок між початковими моментами і твірною функцією початкових моментів

$$\alpha_n = \int_{-\infty}^{\infty} y^n P_\xi(y) dy = \frac{d^n\varphi_\xi(0)}{dx^n}.$$

За означенням твірна функція центральних моментів випадкової величини має вигляд

$$\Phi_{\xi}(x)=\sum_{i=0}^{\infty}\exp[x(y_i-\alpha_1)]P_i$$

для дискретної випадкової величини,

$$\Phi_{\xi}(x)=\int_{-\infty}^{\infty}\exp[x(y-\alpha_1)]P_{\xi}(y)dy$$

для неперервної випадкової величини, де α_1 - перший початковий момент випадкової величини.

Якщо експоненту розвинути в ряд Маклорена

$$\exp[x(y-\alpha_1)]=1+(y-\alpha_1)x+(y-\alpha_1)^2\frac{x^2}{2!}+...+(y-a_1)^n\frac{x^n}{n!}+... ,$$

то для твірної функції моментів отримаємо аналогічне розвинення

$$\Phi_{\xi}(x)=1+\mu_1 x+\mu_2\frac{x^2}{2!}+...+\mu_n\frac{x^n}{n!}+... .$$

Порівнюючи останній ряд з рядом Маклорена для функції $\Phi_{\xi}(x)$

$$\Phi_{\xi}(x)=1+\frac{d\Phi_{\xi}(0)}{dx}x+\frac{d^2\Phi_{\xi}(0)}{dx^2}\frac{x^2}{2!}+...+\frac{d^n\Phi_{\xi}(0)}{dx^n}\frac{x^n}{n!}+...,$$

Отримаємо диференційний зв’язок між центральними моментами і твірною функцією центральних моментів

$$\mu_n=\int_{-\infty}^{\infty}(y-\alpha_1)^n P_{\xi}(y)dy=\frac{d^n\Phi_{\xi}(0)}{dx^n}.$$

Твірну функцію початкових моментів можна задати і в іншому вигляді

$$g_{\xi}(x)=\sum_{i=0}^{\infty}x^i P_i$$

для дискретної випадкової величини,

$$g_\xi(x) = \int_{-\infty}^{\infty} x^y P_\xi(y)dy$$

для неперервної випадкової величини. Таку функцію прийнято називати твірною функцією ймовірності. Легко бачити, що

$$\alpha_1 = \int_{-\infty}^{\infty} yP_\xi(y)dy = \frac{dg_\xi(1)}{dx},$$

$$\alpha_2 = \int_{-\infty}^{\infty} y^2 P_\xi(y)dy = \frac{d^2 g_\xi(1)}{dx^2} + \frac{dg_\xi(1)}{dx},$$

$$\alpha_3 = \int_{-\infty}^{\infty} y^3 P_\xi(y)dy = \frac{d^3 g_\xi(1)}{dx^3} + 3\frac{d^2 g_\xi(1)}{dx^2} + \frac{dg_\xi(1)}{dx},$$

…………………………………………………….. .

Твірну функцію ймовірності для центральних моментів задається так

$$G_\xi(x) = \sum_{i=0}^{\infty} (x-\alpha_1)^i P_i$$

для дискретної випадкової величини,

$$G_\xi(x) = \int_{-\infty}^{\infty} (x-\alpha_1)^y P_\xi(y)dy$$

для неперервної випадкової величини. Легко бачити, що

$$\mu_1 = \int_{-\infty}^{\infty} (y-\alpha_1)P_\xi(y)dy = \frac{dG_\xi(1+\alpha_1)}{dx},$$

$$\mu_2 = \int_{-\infty}^{\infty} (y-\alpha_1)^2 P_\xi(y)dy = \frac{d^2 G_\xi(1+\alpha_1)}{dx^2} + \frac{dG_\xi(1+\alpha_1)}{dx},$$

$$\mu_3 = \int_{-\infty}^{\infty} (y-\alpha_1)^3 P_\xi(y)dy = \frac{d^3 G_\xi(1+\alpha_1)}{dx^3} + 3\frac{d^2 G_\xi(1+\alpha_1)}{dx^2} + \frac{dG_\xi(1+\alpha_1)}{dx},$$

…………………………………………………….. .

Існують і інші варіанти представлення твірної функції моментів

2.8.Закон великих чисел

Закон великих чисел описує поведінку середнього арифметичного n випадкових величин при $n \to \infty$. Він також пов'язує граничне значення цього середнього арифметичного з математичним очікуванням кожної з випадкових величин. Закон великих чисел має декілька формулювань. Конкретне його формулювання залежить від наявної інформації про кожну з випадкових величин $\xi_1, \xi_2, ..., \xi_n$. Фактично закон великих чисел моделює ситуацію з вимірюванням тої чи іншої величини ξ, де $\xi_1, \xi_2, ..., \xi_n$ є результатами окремих вимірювань цієї величини. Закон обґрунтовує і деталізує той факт, що збільшуючи кількість вимірювань (число n) ми, як правило, поліпшуємо, уточнюємо інформацію про невідому величину ξ. Він стверджує, що в міру зростання n середнє арифметичне результатів окремих вимірювань перестає бути випадковою величиною і у межі $n \to \infty$, за певних умов, дає нам точне значення вимірюваної величини. Характерною ознакою закону великих чисел є аналіз не самих відхилень певної випадкової величини від її граничного значення, а ймовірностей цих відхилень. Це цілком природно для випадкової величини, оскільки зазначені відхилення можуть бути довільними, але ймовірність малих відхилень – велика, а великих відхилень – мала.

Найпростіший варіант закону великих чисел виникає тоді, коли всі випадкові величини $\xi_1, \xi_2, ..., \xi_n$ однаково розподілені, попарно незалежні і мають скінчені дисперсії, тоді для довільного $\varepsilon > 0$

$$\lim_{n \to \infty} P\left\{ \left| \overline{\xi} - a \right| < \varepsilon \right\} = 1 .$$

Тут

$$\overline{\xi} = \frac{1}{n}\sum_{k=1}^{n}\xi_k$$

– середнє арифметичне всіх випадкових величин, $M\xi_k = a$.

Величина $\overline{\xi}$ ще називається вибірковим середнім. Даний варіант закону великих чисел можна сформулювати і так. За зазначених вище умов із зростанням об'єму вибірки n вибіркове середнє прямує до математичного очікування випадкової величини ξ, як до своєї межі, тобто

$$M\xi = \lim_{n\to\infty}\overline{\xi} .$$

Складніший варіант закону великих чисел виникає, якщо всі випадкові величини попарно незалежні, але мають свій конкретний закон розподілу з математичним очікуванням $M\xi_k$ і дисперсією $D\xi_k$. Тоді для довільного $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\overline{\xi} - \overline{M\xi}\right| < \varepsilon\right\} = 1 ,$$

де $\overline{\xi}$ – середнє арифметичне випадкових величин, а $\overline{M\xi}$ – середнє арифметичне їх математичних очікувань

$$\overline{M\xi} = \frac{1}{n}\sum_{k=1}^{n} M\xi_k .$$

Закон великих чисел аналізує ситуацію для $n \to \infty$. Як бути, коли n скінчене? Як оцінити ймовірність відхилень значень випадкової величини від свого математичного очікування? Відповідь на ці запитання можна отримати за допомогою відповідної властивості дисперсії, а саме

$$P\left(\left|\xi - M\xi\right| \ge \varepsilon\right) \le \frac{D\xi}{\varepsilon^2} .$$

Використаємо цю нерівність для оцінки наближеного значення вимірюваної величини. Нехай виконане n незалежних вимірювань певної невідомої величини a. Похибки вимірювань $\delta_1, \delta_2, ..., \delta_n$ будемо вважати випадковими величинами. Припустимо, що $M\delta_k = 0$, $1 \le k \le n$. Цю умову можна розглядати як умову відсутності систематичної похибки. Якщо за значення невідомої величини a прийняти середнє арифметичне результатів окремих вимірювань, то так само слід оцінити і похибку у визначенні числа a, а саме:

$$\eta = \frac{\delta_1 + \delta_2 + ... + \delta_n}{n}.$$

При цьому

$$M\eta = M\frac{1}{n}\sum_{k=1}^{n}\delta_k = \frac{1}{n}\sum_{k=1}^{n}M\delta_k = 0,$$

$$D\eta = D\frac{1}{n}\sum_{k=1}^{n}\delta_k = \frac{1}{n^2}\sum_{k=1}^{n}D\delta_k = \frac{\sigma^2}{n},$$

де

$$D\delta_k = \sigma^2.$$

Використовувана нерівність дозволяє оцінити кількість вимірювань потрібно виконати, щоб з великою ймовірністю похибка не перевищувала певне значення ε. Дійсно,

$$P\left(|\eta| \ge \varepsilon\right) \le \frac{D\eta}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Якщо

$$\frac{\sigma^2}{n\varepsilon^2} \le 0.01,$$

то ймовірність того, що похибка перевищує ε не перевищуватиме 0.01. Якщо нас задовольняє така ймовірність, то кількість необхідних для цього вимірювань знаходиться з останньої нерівності, а саме:

$$n \geq 100 \frac{\sigma^2}{\varepsilon^2}.$$

2.9. Центральна межова теорема

Якщо закон великих чисел визначає межове значення вибіркового середнього, то центральна межова теорема визначає поведінку вибіркового середнього, як випадкової величини, в околі цього межового значення. Варіантів центральної межової теореми багато, але головна ідея всіх цих варіантів полягає в тому, що в околі межового значення для вибіркового середнього його поведінка стає універсальною. Тобто, незалежно від того, за яким законом були розподілені окремі доданки суми, аби дисперсія була скінченою, їх середнє арифметичне буде розподілене за нормальним законом. Наведемо одне з формулювань центральної межової теореми.

Теорема. Якщо випадкові величини ξ_1, ξ_2, … , ξ_n незалежні, однаково розподілені і мають скінчену дисперсію, то для $n \to \infty$ рівномірно за x

$$F_\xi(x) = P(\xi < x) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-y^2/2)dy$$

функцією розподілу випадкової величини $\xi = \dfrac{\xi_1 + \xi_2 + ... + \xi_n - na}{\sigma\sqrt{n}}$

буде стандартний нормальний розподіл. Тут $a = M(\xi_1) = M(\xi_2) = ... = M(\xi_n)$ - математичне очікування кожної з випадкових величин, $\sigma^2 = D(\xi_1) = D(\xi_2) = ... = D(\xi_n)$ - дисперсія кожної з випадкових величин. Випадкова величина $\xi = \left(\dfrac{\xi_1 + \xi_2 + ... + \xi_n}{n} - a\right)\dfrac{1}{\sigma/\sqrt{n}}$ - це відхилення вибіркового середнього від математичного очікування, виміряне в одиницях середньоквадратичного відхилення для тої самої вибірки. Розглянута випадкова величина має нульове математичне очікування і одиничну дисперсію. Дійсно,

$$M(\xi) = M\left(\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - a\right)\frac{1}{\sigma/\sqrt{n}}\right) =$$

$$= \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n}M(\xi_i) - M(na)\right) = \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n}a - na\right) = 0\,,$$

$$D(\xi) = D\left(\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - a\right)\frac{1}{\sigma/\sqrt{n}}\right) =$$

$$= M\left(\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - a\right)^2\frac{1}{\sigma^2/n}\right) - \left[M\left(\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - a\right)\frac{1}{\sigma/\sqrt{n}}\right)\right]^2 =$$

$$= \frac{1}{\sigma^2/n}\left(\frac{1}{n^2}\sum_{i,j=1(i\neq j)}^{n}M(\xi_i\xi_j) + \frac{1}{n^2}\sum_{i=1}^{n}M(\xi_i^2) - a^2\right) =$$

$$= \frac{1}{\sigma^2/n}\left(\frac{n(n-1)}{n^2}a^2 + \frac{1}{n^2}\sum_{i=1}^{n}M(\xi_i^2) - a^2\right) = \frac{1}{\sigma^2/n}\left(\frac{1}{n^2}\sum_{i=1}^{n}M(\xi_i^2) - \frac{1}{n}a^2\right) =$$

$$= \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n}\left[M(\xi_i^2) - [M(\xi)]^2\right]\right) = \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n}D(\xi_i)\right) = 1\,.$$

Центральну межову теорему можна записати і у варіанті для густини розподілу для випадкової величини з нульовим математичним очікуванням і одиничною дисперсією

$$P_{\xi}(x) \to \frac{1}{\sqrt{2\pi}}\exp(-x^2/2)\,,$$

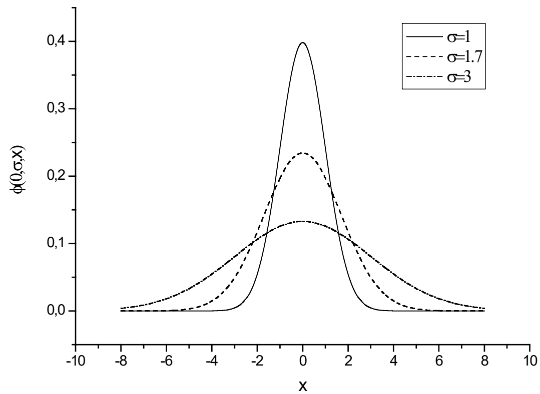
або безпосередньо для вибіркового середнього

$$P_{\bar{\xi}}(x) \to \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\left(\bar{\xi} - a\right)^2}{2\sigma^2/n}\right).$$

Тут $\overline{\xi} = \frac{1}{n}\sum_{i=1}^{n}\xi_i$ - вибіркове середнє. Можна записати також і інтегральний варіант теореми для вибіркового середнього

$$F_{\overline{\xi}}(x) \to \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\overline{\xi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2/n}\right)dx .$$

Нормальний розподіл ймовірності у математичній статистиці займає центральне місце, тому наведемо графік його густини

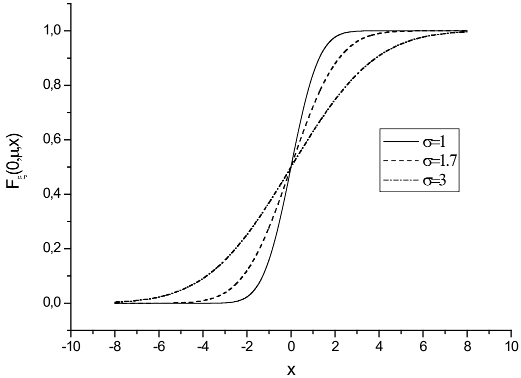


Мал. 1. *Густина нормального розподілу для різних значень дисперсії.*

В загальному випадку функцію нормального розподілу ймовірностей можна записати так

$$F_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} \exp[-(y-\mu)^2/2\sigma^2]dy .$$

Графік цієї функції наступний



Мал. 2. *Функція нормального розподілу в залежності від дисперсії*

Нормальність закону розподілу для випадкової величини $x_m = \left(m - np\right) / \sqrt{npq}$ для достатньо великих, n – кількості випробувань в схемі Бернулі, є одним можливим підґрунтям для центральної межової теореми. Фактично ця теорема збігається з інтегральною або локальною теоремами Муавра-Лапласа. Дійсно, для схеми Бернулі випадковою величиною можна вважати результат окремого випробування. Успіху можна приписати значення 1, невдачі – 0. Закон розподілу такої випадкової величини задано наступною таблицею.

ξ	1	0
P	p	q

Вимірювання цієї випадкової величини полягає в послідовному підкиданні монети і фіксації результату в числовому форматі. Обсягом вибірки буде кількість підкидань. Вибіркове середнє матиме вигляд

$$\overline{\xi} = \frac{1}{n}\sum_{k=1}^{n}\xi_k = \frac{m}{n}.$$

Математичне очікування випадкової величини $M\xi = p$, дисперсія - $D\xi = pq$, вибіркова дисперсія дорівнюватиме $D\overline{\xi} = D\xi / n = pq / n$. Якщо обсяг вибірки достатньо великий, а ймовірність успіху не надто наближається до 1, або 0, то ми отримуємо в якості закону розподілу ймовірностей для вибіркового середнього якраз нормальний закон у повній відповідності з центральною межовою теоремою, але як результат строгої асимптотичної оцінки формули Бернулі.

2.10.Задачі для самостійного розв`язання

Варіант 1.

1.На кожні 40 відштампованих виробів у середньому припадає 4 дефектних. Із усієї продукції навмання узято 400 виробів. Знайти ймовірність того, що серед них 350 виробів будуть без дефектів.

2.Круг рулетки поділено на 12 однакових секторів. Знайти ймовірність того, що у 600 випробуваннях 1-й сектор з'явиться: а) 58 разів; б) не менше, ніж 35 разів; в) від 32 до 55 разів?

3.Ймовірність події дорівнює 0.3. Проведені 200 випробувань. Яка ймовірність того, що частота появи події відрізнятиметься від імовірності не більше, ніж на 0.1? Проведені 200 випробувань. Яким має бути максимальне відхилення частоти від імовірності, щоб ймовірність цього відхилення дорівнювала 0.8?

4. В колоді 36 карт. Яка кількість інформації міститься в повідомленні:

А - чотири шістки розташовані поряд;

В - місця розташування шісток утворюють арифметичну прогресію з кроком 6.

5.Задана випадкова величина

ξ	-2	-1.5	-0.5	1	2	2.5

P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 2.

1.Завод відправив на базу 1000 доброякісних виробів. За час перебування у дорозі кожний виріб було пошкоджено з імовірністю 0.003. Знайти ймовірність того, що на базу прибудуть 3 пошкоджені вироби.

2.У середньому 30% акцій видавничих фірм стають збитковими. Яка ймовірність того, що серед 140 акцій цих фірм збитковими будуть менш як 40%.

3.Ймовірність влучення в ціль при одному пострілі 0.4. Зроблені 200 пострілів. Яка ймовірність того, що частота відрізняється від ймовірності не більше, ніж на 0.04? Скільки потрібно зробити пострілів, щоб з імовірністю 0.96 відхилення частоти від імовірності не перевищувало 0.1?

4.На полиці у випадковому порядку розташовані 40 книжок, серед яких знаходиться тритомник Т.Г. Шевченка. Скільки інформації міститься в повідомленні, що ці томи розташовані у порядку зростання їх номерів але не обов'язково поряд.

5.Задана випадкова величина

ξ	-2.5	-0.5	0	1.5	2	3
P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 3.

1.Для розвинутих країн Заходу частка тіньового бізнесу становить 1%. Яка ймовірність того, що серед 200 зареєстрованих за рік фірм таким бізнесом займаються хоча б дві?

2.Верстат автомат виготовляє стандартну деталь з ймовірністю 0.9. Із продукції беруть партію деталей. Скільки деталей має містити партія, щоб із ймовірністю 0.9973 можна було стверджувати: у партії відхилення відносної частоти появи нестандартної деталі від ймовірності її виготовлення не перевищуватиме 0.03? Визначити можливу кількість нестандартних деталей у партії за даних умов.

3.Ймовірність події дорівнює 0.4. Проведені 400 випробувань. Яка ймовірність того, що частота появи події відрізняється від ймовірності не більше, ніж на 0.06? Проведені 400 випробувань. Яким має бути максимальне відхилення частоти від імовірності, щоб імовірність цього відхилення дорівнювала 0.9?

4. Підкинуті три монети. Скільки інформації міститься в повідомленні:

A - третя монета випала гербом догори;

B - випав рівно один герб;

C - випало рівно два герба;

D - випало не більше двох гербів.

5.Задана випадкова величина

ξ	-4	-2	-0.5	2	3	4.5
P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 4.

1.Зерна пшениці проростають з імовірністю 0.95. Знайти ймовірність того, що із 2000 посіяних зерен зійде від 1880 до 1920.

2.У середньому 80% студентів курсу здають залік з першої спроби. Знайти ймовірність того, що з п'яти навмання взятих студентів цього курсу з першого разу здадуть не більше як четверо.

3.Ймовірність появи події в кожному незалежному випробуванні дорівнює 0.8. Скільки потрібно провести випробувань, щоб з імовірністю 0.95 можна було очікувати відхилення відносної частоти появи від її ймовірності за абсолютної величиною не перевищує 0.03?

4.Підкидається гральний кубик. Скільки інформації міститься в повідомленні:

A - випало число, що не перевищує 2;

B - випало число, що не перевищує 3.

Знайти суму, різницю, добуток і заперечення цих подій.

5.Задана випадкова величина

ξ	-4	-3.5	-0.5	2	3	5
P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 5.

1.Ймовірність розпаду радіоактивного атому протягом року дорівнює 0.0001 . Яка ймовірність того, що з 3000 0 атомів за рік розпадуться: а) 2 атоми; б) не більше 2 атомів; в) хоча б один атом?

2.Відомо, що серед готівкової маси 0.5% купюр непридатні до подальшого використання. Знайти ймовірність того, що серед 2700 купюр виторгу магазину непридатними для наступного використання є хоча б дві купюри?

3.Ймовірність появи події в кожному із 300 незалежних випробувань дорівнює 0.9. Знайти таке додатне число ε, щоб з імовірністю 0.98 абсолютна величина відхилення частоти події від її ймовірності 0.9 не перевищувала ε .

4. В колоді 52 карт. Скільки інформації міститься в повідомленні:

А - чотири шістки розташовані поряд;

В - місця розташування шісток утворюють арифметичну прогресію з кроком 5.

5.Задана випадкова величина

ξ	-7	-4.5	-1.5	1	3	5.5
P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 6.

1.Ймовірність відмови резистора протягом року дорівнює 0.001. В схемі 2000 однакових резисторів. Яка ймовірність того, що в цій схемі відмовить: а) 4 резистори; б) не більше 2 резисторів; в) хоча б один резистор?

2.В урні 2 білі, та три чорні кулі. З урни виймають одну кулю, помічають її колір та повертають кулю назад в урну. Це випробування проводиться 200 разів. Яка ймовірність того, що: а) біла куля з'явиться 70 разів; б) біла куля з'явиться більше 100 разів; в) біла куля з'явиться від 75 до 95 разів; г) число появ білої кулі відрізняється від найбільш ймовірного його числа не більше ніж на 16?

3.Ймовірність події дорівнює 0.4. Проведені 300 випробувань. Яка ймовірність того, що частота появи події відрізнятиметься від імовірності не більше, ніж на 0.2? Проведені 200 випробувань. Яким повинне бути максимальне відхилення частоти від імовірності , щоб імовірність цього відхилення дорівнювала 0.9?

4. Підкинуті три монети. Скільки інформації міститься в повідомленні:

A - всі три монети випали гербом догори;

B - перша монета випала гербом догори;

C - випало рівно два герба;

D - випало не більше одного герба.

5.Задана випадкова величина

ξ	-6	-3.5	-2.5	0.5	2	4.5
P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 7.

1.Відомо, що ймовірність влучення в ціль при одному пострілі дорівнює 0.21. Спортовець зробив 300 пострілів. Яка ймовірність того, що число влучень в ціль: а) дорівнює 55; б) більше, ніж 70; в) знаходиться в межах від 60 до 75; г) відрізняється від максимального значення не більше, ніж на 15?

2.Ймовірність того, що працівник типографії припустився помилки на сторінці тексту, дорівнює 0.005. Яка ймовірність того, що на 600 сторінок: а) ´5 помилок; б) є не більше 2 помилок; в) є хоча б одна помилка?

3.Ймовірність влучення в ціль при одному пострілі 0.3. Зроблені 300 пострілів. Яка ймовірність того, що частота відрізняється від ймовірності не більше, ніж на 0.05? Скільки потрібно зробити пострілів, щоб з імовірністю 0.94 відхилення частоти від імовірності не перевищувало 0.15?

4. Скільки інформації міститься в повідомленні про те, що чотиризначний номер випадково взятого автомобіля:

A - має всі різні цифри;

B - має лише дві однакові цифри;

C - має дві пари однакових цифр;

D -має лише три однакові цифри;

E - має всі однакові цифри.

5.Задана випадкова величина

ξ	-1	-0.5	-0.1	0.1	1.5	2.5
P	0.05	0.2	0.2	0.4	0.05	0.1

Знайти її математичне очікування і дисперсію.

Варіант 8.

1.В лотереї розігрується багато призів. Імовірність виграти будь-який з них дорівнює 0.002. Куплено 2000 лотерейних квитків. Яка ймовірність того, що: а) будуть виграні 6 призів; б) буде виграно не більше 2 призів; в) буде виграно хоча б один приз?

2.Монету підкинули 400 разів. Знайти ймовірність того, що орел з'явиться: а) 210 разів; б) менше, ніж 180 разів; в) не менше 190 та не більше 225 разів; г) число появ орла відрізняється від найбільш ймовірного його числа не більше, ніж на 15?

3.Ймовірність події дорівнює 0.3. Проведені 300 випробувань. Яка ймовірність того, що частота появи події відрізняється від ймовірності не більше, ніж на 0.08? Проведені 200 випробувань. Яким має бути максимальне відхилення частоти від імовірності, щоб імовірність цього відхилення дорівнювала 0.8?

4.В записаному телефонному номері стерлись три останні цифри. Скільки інформації міститься в повідомленні:

A - цифри, що стерлись, різні але відмінні від цифр 1, 2, 3;

B - стерлись однакові цифри;

C - дві цифри, що стерлись, збігаються.

5.Задана випадкова величина

ξ	-2	-1.5	-0.5	1	2	2.5
P	0.04	0.3	0.1	0.4	0.06	0.1

Знайти її математичне очікування і дисперсію.

Варіант 9.

1.З колоди карт виймають одну карту й помічають її масть, після чого карту повертають в колоду. Яка ймовірність того, що в 500 випробуваннях: а) трефа з'явиться 105 разів; б) трефа з'явиться більше, ніж 130 разів; в) трефа з'явиться від 112 до 140 разів; г) число появи масті трефи відхиляється від її максимального значення не більше, ніж на 15 разів?

2.Ймовірність того, що майстер обробить деталь якісно, дорівнює 0.98. Яка ймовірність того, що з 300 оброблених деталей якісно оброблені: а) 295 деталей; б) не менше, ніж 298 деталей; в) не всі деталі якісно оброблені?

3.Ймовірність появи події в кожному незалежному випробуванні дорівнює 0.9. Скільки потрібно провести випробувань, щоб з імовірністю 0.9 можна було очікувати відхилення відносної частоти появи від її ймовірності за абсолютної величиною не перевищує 0.04?

4. На кожній з п'яти карток написані літери Т, А, Р, А, С. Яка ймовірність того, що викладаючи їх в ряд навмання вдасться написати ім’я ТАРАС? Скільки інформації міститься в повідомленні, про те, що навмання вдалося скласти слово Тарас?

5.Задана випадкова величина

ξ	-2	-1.5	-0.5	1	2	2.5
P	0.15	0.1	0.2	0.3	0.05	0.2

Знайти її математичне очікування і дисперсію.

Варіант 10.

1.Ймовірність того, що електричний прилад протягом року працюватиме справно дорівнює 0.97. Яка ймовірність того, що з 200 однакових приладів на підприємстві протягом року справно працюватимуть 196 приладів; б) справно працюватимуть не менше, ніж 198 приладів; в) не всі прилади працюватимуть справно?

2.В урні є 2 білі та 8 чорних куль. З урни виймають одну кулю, помічають її колір і повертають назад в урну. Це випробування провадять 500 разів. Як ймовірність, що: а) біла куля з'явиться 85 разів; б) біла куля з'явиться більше, ніж 120 разів; в) біла куля з'явиться від 75 до 95 разів; г) число появи білої кулі відрізняється від найімовірнішого його значення не більше, ніж на 14?

3.Ймовірність появи події в кожному із 200 незалежних випробувань дорівнює 0.95. Знайти таке додатне число ε, щоб з імовірністю 0.93 абсолютна величина відхилення частоти події від її ймовірності 0.8 не перевищувала ε.

4. На кожній з чотирнадцяти карток написані літери Е, Л, Е, К, Т, Р, И, Ф, І, К, А, Ц, І, Я. Скільки інформації міститься в повідомленні, що навмання вдалося написати слово ЕЛЕКТРИФІКАЦІЯ?

5.Задана випадкова величина

ξ	-2	-1.5	-0.5	1	2	2.5
P	0.2	0.1	0.2	0.3	0.05	0.15

Знайти її математичне очікування і дисперсію.

Глава 3. Елементи математичної статистики

3.1.Вибірка

Інформацію про навколишній світ ми отримуємо в результаті вимірювань. Об'єктом вимірювань може бути і людська спільнота, і тваринний або рослинний світи, і нежива природа. Вимірювана характеристика a часто може мати цілком конкретне числове значення, проте між нами і вимірюваною характеристикою завжди присутній вимірювальний прилад, або вимірювальний алгоритм. Останній, звичайно, при кожному наступному вимірюванні зазначеної характеристики, виконаному за тих самих зовнішніх умов, подає інше її значення. Послідовність результатів вимірювання виглядає як набір значень деякої випадкової величини ξ. Слід розуміти, що цю випадкову величину породжує саме вимірювальний прилад. Досліджуючи таку випадкову величину, ми, фактично, досліджуємо властивості вимірювального приладу. Проте головною задачею математичної статистики є встановлення зв'язку між цією випадковою величиною і вимірюваною характеристикою. В якості основи такого зв'язку постулюється рівність математичного очікування випадкової величини і вимірюваної характеристики

$$M\xi = a .$$

Ця рівність справедлива, так би мовити для ідеального вимірювального приладу. Кожний реальний прилад породжує випадкову величину ξ, математичне очікування якої $M\xi = a^*$ і $a^* \neq a$. Жодним чином математична статистика не здатна ліквідувати розбіжність між a^* і a, але здатна звести до мінімуму вплив на результати вимірювання неконтрольованої дії різноманітних факторів на вимірювальний прилад, тобто зменшити розкиданість значень результатів вимірювань щодо їх математичного очікування $M\xi = a^*$ і оцінити ймовірність можливої відмінності вибіркового середнього і математичного очікування. Ця відмінність виникає через скінченність вибірки, створеної результатами з результатів окремих вимірювань.

Як відомо, для знаходження математичного очікування випадкової величини, потрібно знати для дискретної випадкової величини набір її значень і ймовірностей цих значень, для неперервної випадкової величини – її функцію розподілу ймовірностей, або густину розподілу ймовірностей. Єдиним джерелом такої інформації є результати вимірювань.

Адекватною математичною моделлю вимірюваної величини є, так звана, генеральна сукупність – набір всіх можливих значень вимірюваної величини. Одиницею генеральної сукупності є можливе значення вимірюваної величини. Такий набір можливих значень може утворювати дискретну скінчену, дискретну злічену, неперервну незлічену множину.

Розглянемо приклад скінченої генеральної сукупності. Нехай досліджуються певні характеристики, наприклад середній зріст, певного народу. Кожного окремого представника даного народу можна розглядати як окремий елемент генеральної сукупності. Очевидно у даному разі кількість таких елементів скінчена. Якщо вважати, що зріст кожної людини має цілком конкретне значення, то існує і теоретична, і практична можливість цілком точного обчислення середнього зросту народу. Тоді не виникає потреби у розгляді якоїсь випадкової величини. Потреба в ній з’являється тоді, коли недоцільно проводити глобальне дослідження питання середнього зросту народу, на основи вимірювань зросту всіх його представників. Це може бути технічно складно і дорого. Виникає потреба замінити вихідну генеральну сукупність простішою математичною моделлю, в рамках якої досліджувану характеристику генеральної сукупності можна обчислити хоча і наближено, але з наперед заданою точністю. Такою математичною моделлю є вибірка з генеральної сукупності. Вибірка - це довільна сукупність елементів генеральної сукупності. В нашому прикладі це довільна порівняно невелика група людей вибрана випадковим чином з представників даного народу. Тепер замість генеральної сукупності ми досліджуватимемо вибірку з неї. Середній зріст, обчислений для вибірки, вже буде випадковою величиною через випадковий характер утворення самої вибірки.

Прикладом неперервної генеральної сукупності є, наприклад, геометричні параметри довільного об'єкту: довжина, площа, об'єм тощо. Хоча об'єкт може мати цілком конкретні значення цих параметрів, але дізнатись про них ми можемо лише за допомогою вимірювального приладу. Результати вимірювань можна розглядати як вибірку з незліченої генеральної сукупності, оскільки ми можемо продовжувати вимірювання нескінченно довго, необмежено збільшуючи обсяг вибірки, і ніколи не вичерпаємо всіх її елементів.

Таким чином, нехай потрібно виміряти певну величину a. Результати вимірювань цієї величини $x_1, x_2, ..., x_n$ можна розглядати як набір значень певної випадкової величини ξ, пов'язаної з a, такої що $\xi \in [x_{\min}, x_{\max}]$, де $x_{\min}$ і $x_{\max}$ - мінімальне і максимальне значення випадкової величини ξ. Зазначимо, що така випадкова величина ξ ще називається оцінкою вимірюваної величини a. Кількість різних оцінок, які можна поставити у відповідність вимірюваній величині a нічим не обмежена. Для подальшої розбудови математичного апарату математичної статистики доцільно розглядати кожну з величин $x_1, x_2, ..., x_n$, у свою чергу, як випадкову величину. Тоді замість вибірки $x_1, x_2, ..., x_n$ ми розглядаємо n рівноправних і незалежних копій

$\xi_1, \xi_2, ..., \xi_n$ досліджуваної випадкової величини ξ, які в даній конкретній серії n випробувань набувають значень $x_1, x_2, ..., x_n$. Зрозуміти це можна ось з яких міркувань. Цілком ясно, що чим більший обсяг вибірки, тим більша інформація щодо вимірюваної величини в нашому розпорядженні. Очевидно, у висліді обробки цієї інформації ми точніше зможемо знайти вимірювану характеристику об'єкту досліджень. З цією метою ми можемо збільшувати або обсяг вибірки n, або кількість самих вибірок фіксованого обсягу. В останньому випадку слід, розглянути вибірки

$$x_{11}, x_{12}, ..., x_{1n},$$

$$x_{21}, x_{22}, ..., x_{2n},$$

$$.....................,$$

$$x_{m1}, x_{m2}, ..., x_{mn},$$

$$.....................$$

Сукупність наведених вибірок можна розглядати, як послідовність випадкових величин $\xi_1, \xi_2, ..., \xi_n$, кожна з яких може приймати велику кількість значень:

$$\xi_1 = \{x_{11}, x_{12}, ..., x_{1n}\},$$

$$\xi_2 = \{x_{21}, x_{22}, ..., x_{2n}\},$$

$$..............................,$$

$$\xi_m = \{x_{m1}, x_{m2}, ..., x_{mn}\},$$

$$..............................$$

Відповідно, кожна з випадкових величин $\xi_1, \xi_2, ..., \xi_n$ має певну функцію розподілу ймовірностей. Якщо всі вимірювання проводяться у приблизно однакових умовах, то ці функції розподілу мають також бути приблизно однаковими.

Вибіркові спостереження дають можливість зробити висновки відносно генеральної сукупності шляхом дослідження лише її частини. Для того, щоб ці спостереження - статистичні оцінки були корисними, вони мають відповідати певним вимогам. Основною такою вимогою – фундаментом математичної статистики є вимога того, щоб випадкова величина ξ була **незміщеною** оцінкою вимірюваної величини a, тобто

$$M\xi = a .$$

На базі цього фундаментального співвідношення між вимірюваною величиною a і її реалізації вимірюваним приладом ξ ґрунтується, фактично, математична статистика.

З іншого боку, оскільки, ці ж результати вимірювань можна розглядати як сукупність n випадкових величин, то можна припустити, що також і

$$M\xi_k = a ,\ k = 1,...,n .$$

Тобто властивість незміщеності оцінки може мати дві форми. Надалі ми використовуватимемо обидва підходи. Еквівалентність обох підходів є певною гіпотезою (ергодичною гіпотезою), перевірка якої у кожному конкретному випадку є окремою складною задачею. Надалі ми вважатимемо цю гіпотезу вірною у всіх розглянутих випадках

Введене вище поняття незміщеності оцінки можна дещо розширити.

Означення. Послідовність оцінок ξ_n, $n = 1, 2,...$ називається асимтотично незміщеною послідовністю оцінок параметра a, якщо

$$\lim_{n\to\infty} M\left(\xi_n - a\right) = 0 ,$$

або

$$\lim_{n\to\infty} \xi_n = a .$$

Оцінка ξ параметру a називається спроможною (консистентною, слушною), якщо вона підпорядковується закону великих чисел, тобто для $n \to \infty$ наближається за ймовірністю до шуканого параметру

$$\lim_{n\to\infty} P(|\,\xi_n - a\,| < \varepsilon) = 1.$$

Спроможність означає, що чим більший обсяг вибірки, тим більшою є ймовірність того, що оцінка ξ знаходиться в ε-околі значення параметру a. За достатньо великого обсягу вибірки приналежність оцінки до ε-околу значення параметру a є практично достовірною подією. Спроможність оцінки свідчить про максимально можливу вірогідність отриманої оцінки, її максимальну надійність.

Оцінка ξ параметру a називається **ефективною**, якщо вона має найменшу дисперсію з усіх незміщених і спроможних оцінок.

Оцінка ξ параметру a називається **достатньою**, якщо вона містить всю інформацію вибірки.

Предметом інтересу може бути і функція розподілу випадкової величини ξ у першому підході до результатів вимірювання $F_\xi\left(x\right)$, або набору функцій розподілу $F_{\xi_1}(x), F_{\xi_2}(x), ..., F_{\xi_n}(x)$ у другому підході. Очевидно ці функції розподілу характеризують вимірювальний прилад і дозволяють оцінити надійність отриманих наближених значень для вимірюваної величини a.

Отже, підсумовуючи вищесказане, випадковою вибіркою об'єму n або просто вибіркою називається послідовність незалежних і однаково розподілених випадкових величин $\xi_1, \xi_2, ..., \xi_n$.

Випадкова вибірка є математичною моделлю незалежних вимірювань, що виконуються в однакових умовах. Вибірка є основним об'єктом дослідження математичної статистики. Всі її теоретичні побудови основані на виборці, яка є єдиним математичним об’єктом, отриманим в результаті вимірювань.

Однією з основних характеристик вибірки є емпірична функція розподілу $F_{\xi}^{*}(x)$, що визначається формулою

$$F_{\xi}^{*}(x)=\frac{\mu_n(x)}{n},$$

де $\mu_n\left(x\right)$ – число значень серед $x_1,x_2,...,x_n$ менших за x. Випадкова природа елементів вибірки $\xi_1,\xi_2,...,\xi_n$ робить випадковою і величину $\mu_n\left(x\right)$.

Крім вибірки $x_1,x_2,...,x_n$ обсягом n, іншим близьким до неї об’єктом є варіаційний ряд. Він отримується з вибірки упорядкуванням її членів у порядку зростання. Отже варіаційний ряд має вигляд

$$x_{(1)}\le x_{(2)}\le ...\le x_{(n)}.$$

Очевидно, що

$$x_{(1)}=\min\left\{x_1,x_2,...,x_n\right\},\quad x_{(n)}=\max\left\{x_1,x_2,...,x_n\right\}$$

тощо Можна довести, що для великих обсягів вибірки емпірична функція розподілу $F_{\xi}^{*}(x)$ близька до функції розподілу $\mathrm{F}_{\xi}\left(x\right)$ або сукупності функцій розподілу $F_{\xi_i}(x)$. Математично це твердження можна записати так: для довільного $x\in\left(-\infty,\infty\right)$ і довільного $\varepsilon>0$

$$\lim_{n\to\infty}P\left\{\left|F_{\xi}^{*}(x)-F_{\xi}(x)\right|<\varepsilon\right\}=1.$$

Для наочного представлення вибірки використовується **гістограма**. Гістограма будується так. Числова пряма розбивається на декілька неперетинних менших інтервалів

$$(-\infty,\infty)=\sum_{k=0}^{n}[z_k,z_{k+1}),$$

де

$$-\infty < z_0 < z_1 < \ldots < z_r < z_{r+1} < \infty .$$

Далі обчислюється частота попадання елементів вибірки в кожний з цих менших інтервалів. Ці частоти P_k^* можна обчислити безпосередньо. Якщо відома емпірична функція розподілу вибірки, то частоти можна визначити за так

$$P_k^* = F_\xi^*(z_{k+1}) - F_\xi^*(z_k) .$$

Останнє співвідношення нагадує відповідне співвідношення для неперервної випадкової величини

$$P\left(z_k \le \xi < z_{k+1}\right) = F_\xi\left(z_{k+1}\right) - F_\xi\left(z_k\right).$$

Можна довести, що для великих обсягів вибірки n останні два співвідношення стають практично ідентичними. Якщо теоретична або експериментальна функція розподілу $F_\xi\left(x\right)$ має густину розподілу $P_\xi\left(x\right)$, то для великих обсягів вибірки

$$P_k^* \approx P\left(z_k \le \xi < z_{k+1}\right) = \int_{z_k}^{z_{k+1}} P_\zeta\left(x\right) dx .$$

Оскільки при відповідному виборі точок z_k

$$\int_{z_k}^{z_{k+1}} P_\xi\left(x\right) dx \approx P_\xi\left(x_k\right) \Delta z_k ,$$

то

$$P_k^* \approx P_\xi\left(x_k\right) \Delta z_k$$

і гістограмам нагадує графік густини розподілу. Очевидно, що чим менші проміжки $\Delta z_k = z_{k+1} - z_k$ ми оберемо і чим менше вони відрізнятимуться між собою, тим зазначена подібність буде більшою.

3.2.Вибіркові моменти

Вибірка, як і кожний математичний об’єкт, має певний набір характеристик. Всі характеристики вибірки називаються вибірковими. До найважливіших вибіркових характеристик належать вибіркові моменти: початкові і центральні. **Початкові вибіркові моменти** визначаються так

$$a_\nu = \frac{1}{n}\sum_{k=1}^{n} x_k^\nu ,$$

подібним чином визначаються і **центральні вибіркові моменти**

$$m_\nu = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{x}\right)^\nu ,$$

де

$$\overline{x} = a_1 = \frac{1}{n}\sum_{k=1}^{n} x_k$$

– перший початковий вибірковий момент або вибіркове математичне очікування, або просто вибіркове середнє. Для вибіркової дисперсії або другого центрального моменту випадкової величини часто використовують позначення

$$s^2 = m_2 = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{x}\right)^2 .$$

Вибіркові моменти і початкові, і центральні є випадковими величинами але ступінь їх випадковості є меншою за ступінь випадковості вихідної випадкової величини і ще зменшується із збільшенням обсягу вибірки.

Теоретичні моменти в теорії ймовірностей можна знайти лише, якщо відомі функція розподілу ймовірностей або густина розподілу ймовірностей у разі неперервної випадкової величини. У разі дискретної випадкової величини слід знати ймовірності різних значень випадкової величини. Вибірка ж має всі ознаки дискретної випадкової величини із скінченим набором значень.

Тобто до неї слід застосувати алгоритм розгляду дискретної випадкової величини. Проте інформації про ймовірності різних її значень у нас немає. Єдине практично виправдане припущення полягає у тому, що всі елементи вибірки, навіть якщо деякі її елементи зустрічаються декілька разів, рівноправні між собою і ймовірності їх появи у вибірці однакові і дорівнюють $P_1 = P_2 = ... = P_n = 1/n$. Саме це припущення і дає для експериментальних початкових і центральних моментів наведені вище вирази.

З іншого боку, генеральна сукупність це набір всіх теоретично можливих значень випадкової величини ξ, породжуваної вимірювальним приладом. Цей набір значень найчастіше є неперервним. Тоді можна говорити про випадкову неперервну величину ξ, а вибірку вважати лише певною частковою інформацією про її можливі значення. У цьому разі можна реалізувати, як це було виконано у попередніх параграфах, алгоритм знаходження її густини розподілу ймовірностей або функції розподілу ймовірностей. Знайшовши останню функцію, ми можемо, аналогічно теоретичним моментам, знайти експериментальні моменти. Виникає запитання, який з цих підходів точніший. Можна довести, що перший, якщо при побудові експериментальної густини розподілу ми розбили проміжок між максимальним і мінімальним елементами вибірки на доволі великі менші проміжки. Ця розбіжність прямуватиме до нуля, якщо довжина цих менших проміжків також прямуватиме до нуля.

Зазначимо, що вся сукупність початкових або центральних вибіркових моментів несе максимально можливу інформацію про вибірку. Іншими словами, сукупність всіх вибіркових моментів містить інформацію, аналогічну емпіричній функції розподілу $F_\xi^*(x)$. Зручність використання вибіркових моментів полягає у тому, що для практичного використання інформації про вибірку досить знати лише перші декілька вибіркових моментів. При досить великому обсязі вибірки вибіркові характеристики a_ν, m_ν бувають близькими до відповідних характеристик теоретично існуючої, але нам невідомої функції $F_\xi(x)$, тобто до її відповідних моментів

$$\alpha_\nu = M\xi^\nu$$

та

$$\mu_\alpha = M\left(\xi - M\xi\right)^\nu .$$

Теоретичні моменти і початкові, і центральні є вже числовими характеристиками випадкової величини і в цьому їх головна зручність.

Проведемо теоретичний аналіз введених нами вибіркових моментів. Оскільки $x_1, x_2, ..., x_n$ за означенням є випадковими величинами, то вибіркові моменти, оскільки вони є певними комбінаціями випадкових величин, теж є випадковими величинами. Знайдемо їх математичне очікування і дисперсію, використовуючи теоретично існуючі функції розподілу $\mathrm{F}_{\xi_1}(x_1), \mathrm{F}_{\xi_2}(x_2), ..., \mathrm{F}_{\xi_n}(x_n)$. За властивістю математичного очікування математичне очікування суми випадкових величин є сумою математичних очікувань окремих випадкових величин. Отже

$$M\overline{x} = M\frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{n}\sum_{k=1}^{n} Mx_k = \frac{1}{n}\sum_{k=1}^{n} a = \alpha_1 .$$

Оскільки випадкові величини не тільки однаково розподілені, але і незалежні, то подібна властивість має місце і для дисперсії. Отже

$$D\overline{x} = D\frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{n^2} D\sum_{k=1}^{n} x_k = \frac{1}{n^2}\sum_{k=1}^{n} Dx_k =$$

$$= \frac{1}{n^2} \sum_{k=1}^{n} \mu_2 = \frac{\mu_2}{n}.$$

Таким чином,

$$M\overline{x} = \alpha_1 \ ,$$

$$D\overline{x} = \frac{\mu_2}{n}.$$

Аналогічно можна знайти, що для $\nu > 1$

$$Ma_\nu = \alpha_\nu ,$$

$$Da_\nu = \frac{\alpha_{2\nu} - \alpha_\nu^2}{n}.$$

Складнішим є обчислення математичного очікування і дисперсії центральних вибіркових моментів, тому наведемо лише результати Mm_2 та Dm_2:

$$Mm_2 = \frac{n-1}{n} \mu_2 ,$$

$$Dm_2 = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - 2\mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}.$$

Як ми бачимо, математичне очікування початкових вибіркових моментів збігається з відповідними теоретичними моментами. Дисперсія початкових вибіркових моментів вже відрізняється від теоретичних моментів за рахунок початкових моментів подвоєного порядку. Математичне очікування другого центрального моменту відрізняється від відповідного теоретичного моменту μ_2 на величину $-\mu_2 / n$. Дисперсія другого центрального моменту взагалі визначається не лише через теоретичний другий момент але і четвертий.

Можна довести, що для великих n

$$Mm_\nu = \mu_\nu + D\left(\frac{1}{n}\right),$$

що добре видно на прикладі другого вибіркового моменту, для якого

$$Mm_2 = \mu_2 - \frac{\mu_2}{n}.$$

3.3.Інтегрування Методом Монте-Карло

Розглянемо обчислення визначеного інтегралу

$$I = \int_a^b f(x)dx.$$

Його можна представити як знаходження математичного очікування випадкової функції – початкового моменту першого порядку з густиною розподілу ймовірності $P_\xi(x)$

$$Mf(\xi) = \int_a^b P_\xi(x) f(x)dx,$$

Якщо в якості густини розподілу взяти густину рівномірного розподілу

$$P_\xi(x) = \frac{1}{b-a}\begin{cases} 0, x < a, \\ 1, a \le x \le b, \\ 0, x > b. \end{cases}$$

Тоді

$$Mf(\xi) = \frac{1}{b-a}\int_a^b f(x)dx$$

або

$$\int_a^b f(x)dx = (b-a)Mf(\xi).$$

В якості точкової оцінки математичного очікування можна взяти вибіркове середнє для рівномірно розподіленої генеральної сукупності

$$Mf(\xi) \approx \frac{1}{n}\sum_{k=1}^{n} f(x_k).$$

Отже,

$$\int_a^b f(x)dx \approx \frac{b-a}{n}\sum_{k=1}^{n} f(x_k).$$

Очевидно, що точність метода тим більша, чим більший обсяг вибірки.

Приклад 1. *Розглянемо реалізацію методу в пакеті Mathcad.*

$$a := 1,\ b := 4,\ f(x) := \exp(-x).$$

Значення інтегралу, обчисленого методом Рунге-Кута

$$\int_a^b f(x)dx = 0.3496.$$

Для обчислення цього ж інтегралу методом Монте-Карло використаємо формулу

$$I(n) := \frac{b-a}{n}\sum_{k=1}^{n} f(a+(b-a)\cdot rnd(1)).$$

Тут функція Mathcad rnd(1) при кожному звертанні до неї видає випадкове число на відрізку $[0,1]$. *У висліді вираз* $a+(b-a)\cdot rnd(1)$ *набуває довільного значення на відрізку* $[a,b]$.

Використання методу Монте-Карла вимагає вибірок великого обсягу. При цьому кількість операцій значно перевищуватиме при тій самій точності кількість операцій при використанні не стохастичних методів. Переваги цього методу стають очевидними при обчисленні багатократних інтегралів і інтегралів від складних функцій. Нижче можна побачити *залежність точності обчислення інтегралу методом Монте-Карло від обсягу вибірки*

$$I(10) = 0.212,\ I(100) = 0.378,\ I(1000) = 0.365,$$

$$I(10000) = 0.347,\ I(10000) = 0.353.$$

У даному прикладі стабілізувати навіть другу значущу цифру доволі важко.

3.4.Вибіркові моменти для великих вибірок

Для оцінки поведінки вибіркових моментів для великих обсягів вибірки $n \to \infty$ важливими є наступні твердження.

Теорема 1. Якщо скінченими є теоретичні моменти парних порядків $\alpha_{2\nu}$, то для $n \to \infty$ вибірковий момент a_ν асимтотично нормальний з параметрами $\left(\alpha_\nu, \dfrac{\alpha_{2\nu} - \alpha_\nu^2}{n}\right)$.

Теорема 2. Якщо скінченим є теоретичний момент μ_4, то для $n \to \infty$ вибіркова дисперсія m_2 асимптотично нормальна з параметрами $\left(\mu_2, \dfrac{\mu_4 - \mu_2^2}{n}\right)$.

Отже, межові функції розподілу вибіркових моментів є нормальними, що є прямим наслідком центральної межової теореми. Однак, домежові функції розподілу різних вибіркових характеристик можуть суттєво відрізнятись від нормальних, тому знаходження функцій розподілу для вибірок скінченого обсягу є надзвичайно складною задачею. При цьому, якщо в рамках обмежень теорем 1 і 2 межові функції розподілу вибіркових моментів не залежать від функцій розподілу випадкових величин, то домежові функції розподілу можуть залежати досить суттєво. Фактично більш-менш простою задачею є знаходження домежових функцій розподілу лише у разі, випадкових величин

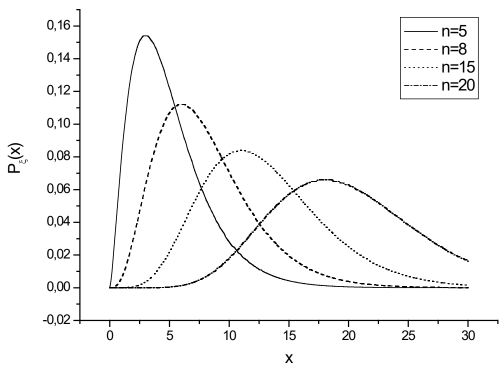
$x_1, x_2, ..., x_n$, що утворюють вибірку, і мають нормальний розподіл. Нагадаємо, що функції розподілу $F_\xi(x)$ або $F_{\xi_1}(x_1), F_{\xi_2}(x_2), ..., F_{\xi_n}(x_n)$ – характеризують вимірювальний прилад. У разі нормально розподілених випадкових величин $\xi_0, \xi_1, ..., \xi_n$ з параметрами $(0,1)$, розподіл випадкової величини

$$\chi_n^2 = \xi_1^2 + \xi_2^2 + ... + \xi_n^2,$$

відомий і називається розподілом χ^2 з n – ступенями вільності. Густина цього розподілу має вигляд

$$\mathrm{P}_{\chi_n^2}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \ x > 0, \ \chi_n^2 = \xi_1^2 + ... + \xi_n^2.$$

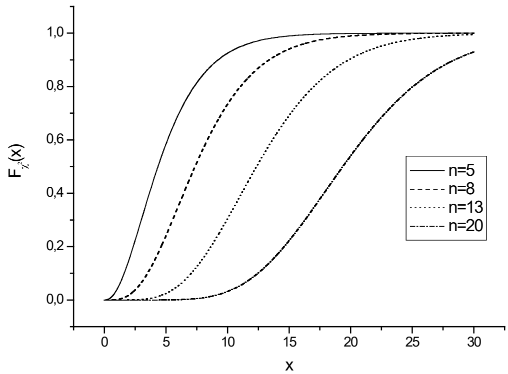
Його графік наступний



Мал. 1. *Густина розподілу* χ_n^2 *для різних значень* n.

Функція розподілу ймовірностей χ^2 має вигляд

$$F_{\chi_n^2}(x) = \int_0^x \mathrm{P}_{\chi_n^2}(x)\,dx$$



Мал. 2. *Функція розподілу* χ_n^2 *- квадрат в залежності від* n.

Для великих обсягів вибірки розподіл χ^2 практично не відрізняється від нормального і практично не залежить від обсягу вибірки. Для малих обсягів вибірки розподіл χ^2 суттєво відрізняється від нормального, а також суттєво залежить від обсягу вибірки.

Розподіл випадкової величини

$$\tau_n = \frac{\xi_0}{\sqrt{\frac{1}{n}\chi_n^2}}$$

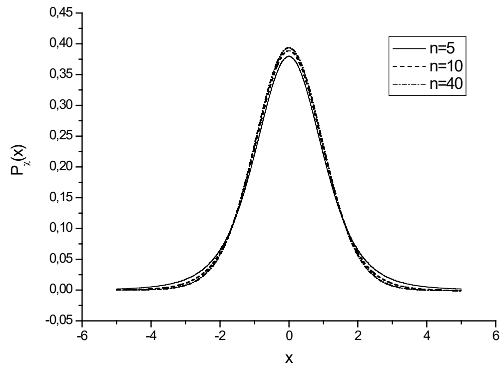
називається розподілом Стьюдента з n ступенями вільності. Густина розподілу Стьюдента має вигляд

$$\mathrm{P}_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \ \tau_n = \frac{\xi_0}{\sqrt{\frac{1}{n}\chi_n^2}}, \quad x \in (-\infty, \infty).$$

Відповідно для функції розподілу матимемо

$$F_{\tau_n}(x) = \int_{-\infty}^{x} \mathrm{P}_{\tau_n}\left(x\right)dx .$$

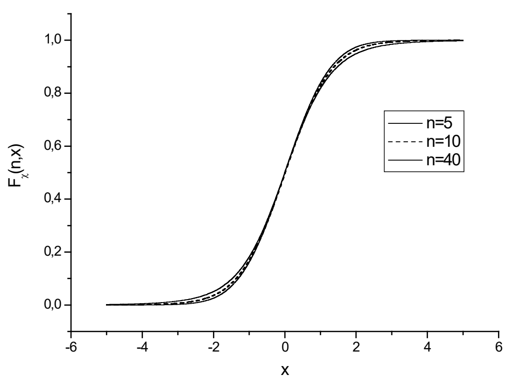
Графік густини розподілу наступний



Мал. 3. *Густина розподілу Стьюдента для різних* n.

Для великих обсягів вибірки розподіл Стьюдента практично не відрізняється від нормального і практично не залежить від обсягу вибірки. Для малих обсягів вибірки розподіл Стьюдента дещо відрізняється від нормального і залежить від обсягу вибірки.

Графік функції розподілу Стьюдента наступний



Мал. 4. Функція розподілу Стьюдента для різних n.

Якщо $\chi^2_{n_1}$ та $\chi^2_{n_2}$ є незалежними випадковими величинами, що мають розподіл χ^2 з n_1 і n_2 ступенями вільності відповідно, то розподіл випадкової величини

$$\mathrm{F}_{n_1,n_2}=\frac{\chi^2_{n_1}/n_1}{\chi^2_{n_2}/n_2}$$

називається F – розподілом або розподілом Фішера з n_1 і n_2 ступенями вільності. Густину розподілу Фішера також можна знайти явно.

Наступні теореми визначають зв'язок між характеристиками вибірки і зазначеними вище розподілами.

Теорема 3. Якщо елементи вибірки $x_1,x_2,...,x_n$ незалежні і розподілені нормально з параметрами $\left(a,\sigma^2\right)$, то вибіркове середнє $\overline{x}$ і вибіркова дисперсія $s^2=m_2$ незалежні, причому $\overline{x}$ розподілене нормально з параметрами $\left(a,\sigma^2/n\right)$, а ns^2/σ^2 має розподіл χ^2 з $n-1$ ступенями вільності.

Теорема 4. Якщо елементи вибірки $x_1,x_2,...,x_n$ незалежні і розподілені нормально з параметрами $\left(a,\sigma^2\right)$, то величина $\dfrac{\overline{x}-a}{s}\sqrt{n-1}$ має розподіл Стьюдента з $n-1$ ступенями вільності.

Теорема 5. Якщо $x_1,...,x_{n_1}$ і $x_1',...,x_{n_2}'$ – дві незалежні вибірки і кожна з n_1+n_2 випадкових величин є незалежною і нормально розподіленою з параметрами $\left(a,\sigma^2\right)$, то випадкова величина $\frac{s^2}{s^{2\,\prime}}\cdot\frac{n_1}{n_2}\cdot\frac{n_2-1}{n_1-1}$, де $s^2=m_2$ – вибіркова дисперсія вибірки $x_1,...,x_{n_1}$, а $s^{2\,\prime}=m_2'$ – вибіркова дисперсія вибірки $x_1',...,x_n'$, мають розподіл Фішера з n_1-1 і n_2-1 степенями свободи. Густина розподілу Фішера має вигляд

$$\mathrm{P}_{F_{n_1,n_2}}(x)=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}\frac{x^{\frac{n}{2}-1}}{\left(\frac{n_1}{n_2}x+1\right)^{\frac{n_1+n_2}{2}}},\ x>0,$$

$$F_{n_1,n_2}=\frac{\chi^2_{n_1}/n_1}{\chi^2_{n_2}/n_2}.$$

Тут всі випадкові величини $\xi_0,\xi_1,...,\xi_n$ розподілені за нормальним законом з параметрами $(0,1)$.

3.5.Емпірична функція розподілу

Розглянемо детальніше емпіричну функцію розподілу. Нагадаємо, що емпіричною функцією розподілу, або функцією розподілу вибірки називають функцію $F_\zeta^*(x)$, яка визначає для кожного значення x частоту події $\zeta<x$

$$\mathrm{F}_\zeta^*(x)=\frac{n_x}{n}.$$

Тут n – обсяг вибірки, n_x – кількість значень (кількість варіантів) випадкової величини ζ, які менше від x. Нагадаємо також, що теоретична функція розподілу визначає ймовірність того, що значення випадкової величини $\zeta<x$, тобто

$$F_{\zeta}(x) = P(\zeta < x).$$

При цьому

$$\lim_{n\to\infty} F_{\zeta}^{*}(x) = F_{\zeta}(x).$$

Для скінчених n

$$F_{\zeta}^{*}(x) \approx F_{\zeta}(x).$$

Емпірична функція розподілу $F_{\zeta}^{*}(x)$ має такі властивості:

1. $0 \le F_{\zeta}^{*}(x) \le 1$;
2. $F_{\zeta}^{*}(x)$ – зростаюча функція;
3. $F_{\zeta}^{*}(x) = \begin{cases} 0, \; x \le x_{\min}; \\ 1, \; x > x_{\max}, \end{cases}$

де $x_{\min}$ – найменше значення елемента вибірки, $x_{\max}$ – найбільше значення елемента вибірки.

Задача 1. *Знайти емпіричну функцію розподілу за статистичним розподілом вибірки*

x_i	2	6	10
n_i	12	18	30

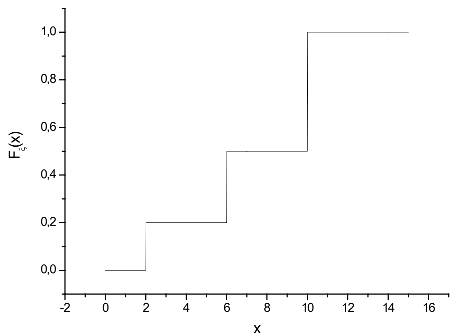
та побудувати її графік.

Розв`язання. *Об'єм цієї вибірки буде* $n = 12 + 18 + 30 = 60$. *Найменше значення елементів вибірки дорівнює 2, тому* $F_{\zeta}^{*}(x) = 0$ *для* $x < 2$. *Найбільше значення елементів вибірки (найбільша варіанта) дорівнює 10, тому* $F^{*}(x) = 1$ *для* $x \ge 10$. *Значення* $\zeta < 6$, *тобто* $\zeta = \{x_1 = 2\}$, *спостерігається 12 разів, тому* $F_{\zeta}^{*}(x) = 12/60 = 0.2$ *при* $2 \le x < 6$. *Значення* $\zeta < 10$, *тобто* $\zeta = \{x_1 = 2\} \cup \{x_2 = 6\}$ *спостерігається 12+18=30 разів, тому*

$F_{\zeta}^{*}(x) = 30/60 = 0.5$ *при* $6 \le x < 10$. *Таким чином, одержали емпіричну функцію розподілу вигляду*

$$F_{\zeta}^{*}(x) = \begin{cases} 0, \; x < 2; \\ 0.2, \; 2 \le x < 6; \\ 0.5, \; 6 \le x < 10; \\ 1, \; x \ge 10. \end{cases}$$

Графік цієї функції наступний



Мал. 1. *Емпірична функція розподілу ймовірностей*

Задача 2. *У висліді спостереження одержали розподіл ознаки* ζ *вибірки у вигляді*

x_k	*-2*	*0*	*1*	*2*	*3*	*5*	*7*
n_k	*4*	*5*	*7*	*8*	*6*	*2*	*1*

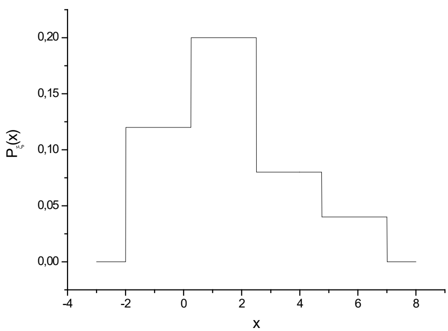
Побудувати гістограму частот цього розподілу.

Розв`язання. *У цьому разі найменше значення випадкової величини* ζ *є* $x_{\min} = -2$, *а найбільше значення -* $x_{\max} = 7$, *тому довжина проміжку* $[x_{\min}, x_{\max}]$ – $L = x_{\max} - x_{\min} = 9$. *Розіб'ємо цей відрізок на* $m = 4$ *рівних частини довжиною* $h = L/m = 9/4 = 2.25$

. Для побудови гістограми доцільно скласти наступну допоміжну таблицю: у перший рядок таблиці

$h = 2.25$	*[-2, 0.25)*	*[0.25, 2.5)*	*[2.5, 4.75)*	*[4.75, 7]*
n_k	*9*	*15*	*6*	*3*
n_k / nh	*0.12*	*0.20*	*0.08*	*0.04*

запишемо одержані відрізки, у другий рядок – кількість елементів вибірки, що потрапляють у відповідні інтервали, у третій – висоти відповідних прямокутників. За даними цієї таблиці будуємо відповідну гістограму



Мал. 2. *Емпірична густина розподілу ймовірностей.*

Гістограмою частот називають ступінчату фігуру, яка складається з прямокутників, основами яких є часткові інтервали довжиною $h = x_k - x_{k-1}$, а висоти дорівнюють густинам частот n_k / nh. Тут n – обсяг вибірки. Очевидно, площа гістограми дорівнює 1.

Отже, для побудови гістограми проміжок, якому належить вибірка $[x_{\min}, x_{\max}]$, де $x_{\min}$ – мінімальне значення елемента вибірки, $x_{\max}$ – максимальне значення елемента вибірки, розбивається на m відрізків $[x_i, x_{i+1})$, $i = 1,...,m$, $x_1 = x_{\min}$, $x_{m+1} = x_{\max}$ рівної довжини h. Потім підраховується кількість елементів вибірки, значення яких належать кожному з інтервалів. Нарешті будуються прямокутники з основою h і висотою n_k / nh.

Медіаною емпіричної густини ймовірностей є точка на горизонтальній осі така, що перпендикулярна лінія, яка проходить через неї, поділяє площу гістограми на дві рівні за площею частини.

3.6. Оцінювання параметрів розподілу

Метою використання вибірки може бути як наближене обчислення тої, чи іншої характеристики об'єкту, так і обчислення параметрів, що визначають функцію розподілу окремих елементів вибірки, які розглядаються як випадкові величини, або різних вибіркових характеристик, які також є випадковими величинами. Нижче ми розглянемо саме останню задачу - задачу оцінювання параметрів розподілів.

Нехай випадковий вектор $\zeta = (\zeta_1, \zeta_2, ..., \zeta_n)$ є вибіркою з генеральної сукупності або вибірковим вектором. Тут $\zeta_1, \zeta_2, ..., \zeta_n$ – незалежні, однаково розподілені випадкові величини. Якщо кожна з цих випадкових величин має розподіл G, то їх сукупність називають вибіркою обсягом n із розподілу G.

Якщо закон розподілу G невідомий, то виникає проблема його знаходження. У найпростішій ситуації закон розподілу G, тобто функція $F_\zeta(x, \theta)$, може бути відомим з точністю до невідомого параметра θ. Тоді виникає задача за даною вибіркою встановити більш або менш точно значення цього параметра. Саме останню задачу ми і розглянемо в даному параграфі.

Єдиним джерелом інформації, від якого залежить можливе значення параметра розподілу θ, є тільки сама вибірка. Отже задача оцінювання параметра розподілу полягає у побудові такої функції реалізації вибірки $h(\xi)$, яка б точно або наближено дорівнювала θ.

Принциповим питанням задачі оцінювання параметрів розподілу є оцінка величини похибки $\hat{\theta}-\theta$, де θ – невідоме точне значення параметра, а $\hat{\theta}$ – відоме наближене. Найпростішою мірою похибки (однією з багатьох можливих), очевидно, є наступна величина $M(\hat{\theta}-\theta)^2$, тобто середнє значення квадрата відхилення випадкової величини $\hat{\theta}$ від її точного значення θ. Можна показати, що серед різних оцінок $\hat{\theta}$ з тією самою дисперсією мінімальна похибка (мінімальне розсіювання) відповідатиме оцінкам, для яких $M\hat{\theta}=\theta$, тобто для незміщених оцінок. Дійсно,

$$M\left(\hat{\theta}-\theta\right)^2 = M\left(\left(\hat{\theta}-M\hat{\theta}\right)+\left(M\hat{\theta}-\theta\right)\right)^2 =$$

$$= M\left(\hat{\theta}-M\hat{\theta}\right)^2 + 2M\left(\hat{\theta}-M\hat{\theta}\right)\left(M\hat{\theta}-\theta\right)+M\left(M\hat{\theta}-\theta\right)^2 =$$

$$= D\hat{\theta}+2\left(M\hat{\theta}-M\hat{\theta}\right)\left(M\hat{\theta}-\theta\right)+\left(M\hat{\theta}-\theta\right)^2 = D\hat{\theta}+\left(M\hat{\theta}-\theta\right)^2.$$

Тут ми врахували, що за означенням дисперсії

$$D\hat{\theta}=M\left(\hat{\theta}-M\hat{\theta}\right)^2$$

і

$$M\theta=\theta,\ MM\hat{\theta}=\hat{\theta}.$$

Отже, якщо $M\hat{\theta}=\theta$, то помилка оцінювання є мінімальною і дорівнює дисперсії. На практиці незміщеність оцінки $\hat{\theta}$ параметра θ можна трактувати так: при багаторазовому використанні оцінки $\hat{\theta}$ як значення θ середнє значення похибки $\hat{\theta}-\theta$ дорівнює нулю.

3.7.Точкові **оцінки**

Нехай елементи вибірки $x_1, x_2, ..., x_n$ залежать від невідомого параметра θ. Очевидно від цього параметра буде залежати функція розподілу кожного елемента вибірки

$$P\left(x_k < x\right) = F_{\xi}\left(x, \theta\right).$$

Оцінкою θ_n^* параметра θ називається довільна функція $\theta_n^* = \theta_n^*\left(x_1, ..., x_n\right)$. Таким чином, θ_n^* є випадковою величиною. Надалі нас цікавитимуть незміщені оцінки θ_n^* параметру θ, тобто такі, що для довільного n

$$M\theta_n^* = \theta,$$

а також слушні оцінки, тобто такі, що для довільного $\varepsilon > 0$ і $n \to \infty$

$$P\left(\left|\theta_n^* - \theta\right| < \varepsilon\right) \to 1.$$

Якщо незміщені і слушними є декілька різних оцінок одного і того ж параметра, то з них природно обрати ефективну оцінку, тобто таку, що має меншу дисперсію. Для кількісної оцінки декількох оцінок одного і того ж параметра за останнім критерієм використовується наступна величина

$$e\left(\theta_n^*\right)=1/\left[nI(\theta)D\theta_n^*\right],$$

причому

$$0\le e\left(\theta_n^*\right)\le 1\,.$$

Якщо $e\left(\theta_n^*\right)=1$, то оцінку називають ефективною. Тут $D\theta_n^*$ – дисперсія оцінки, а

$$I(\theta)=M\left[\frac{\partial\ln P\left(x_k,\theta\right)}{\partial\theta}\right]^2,$$

де $P\left(x_k,\theta\right)$ – густина розподілу елемента вибірки x_k.

Задача 1. *Нехай* $\xi=(x_1,x_2,...,x_n)$ *– вибірка, що описується нормальним законом з параметрами* $\left(a,\sigma^2\right)$, *що є математичним очікуванням і дисперсією окремих елементів вибірки. В якості оцінок* a *і* σ^2 *візьмемо вибіркове середнє*

$$\overline{x}=\frac{1}{n}\left(x_1+x_2+...+x_n\right)$$

і виправлену вибіркову дисперсію

$$S^2=\frac{n}{n-1}m_2=\frac{1}{n-1}\sum_{k=1}^{n}\left(x_k-\overline{x}\right)^2.$$

Доведемо, що обидві оцінки є незміщеними і слушними.

Розв`язання. *Дійсно, оскільки*

$$Mx_1=Mx_2=...=Mx_n=a\,,$$

то

$$M\overline{x}=M\frac{1}{n}\left(x_1+x_2+...+x_n\right)=$$

$$=\frac{1}{n}\left(Mx_1+Mx_2+...+Mx_n\right)=Mx_i=a\,.$$

Відповідно, оскільки,

$$Dx_1 = Dx_2 = \ldots = Dx_n = \sigma^2 ,$$

$$m_2 = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{x}\right)^2 = \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \overline{x}^2 ,$$

$$Mx_k = 0 ,$$

$$Mx_k x_l = Mx_k Mx_l = 0, \ (k \neq l),$$

то

$$M\overline{x}^2 = M\frac{1}{n^2}\sum_{k,l=1}^{n} x_k x_l = \frac{1}{n^2}\sum_{k,l=1}^{n} Mx_k x_l = \frac{1}{n^2}\sum_{k=1}^{n} Mx_k^2 ,$$

$$Mm_2 = M\frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{x}\right)^2 = M\left[\frac{1}{n}\sum_{k=1}^{n} x_k^2 - \overline{x}^2\right] =$$

$$= \frac{1}{n}\sum_{k=1}^{n} Mx_k^2 - \frac{1}{n^2}\sum_{k=1}^{n} Mx_k^2 = \frac{n-1}{n^2}\sum_{k=1}^{n} Mx_k^2$$

$$MS^2 = \frac{n}{n-1}Mm_2 = \frac{1}{n}\sum_{k=1}^{n} Mx_k^2 =$$

$$= \frac{1}{n}\sum_{k=1}^{n}(Mx_k^2 - (Mx_k)^2) = \frac{1}{n}\sum_{k=1}^{n} Dx_k = \sigma^2 .$$

Розглянемо тепер дисперсію цих оцінок

$$D\overline{x} = D\frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{n^2}\sum_{k=1}^{n} Dx_k = \frac{\sigma^2}{n} .$$

Аналогічно можна довести, що

$$DS^2 = \frac{\mu_n - \mu_2^2}{n} + D\left(\frac{1}{n^2}\right).$$

Задача 2. *Нехай* $\zeta_1, \zeta_2, \ldots, \zeta_n$ *- вибірка з розподілу з густиною*

$$f\left(x,a,b\right) = \frac{1}{b-a}\begin{cases} 1, \ if \ x \in [a,b], \\ 0, \ if \ x \notin [a,b]. \end{cases}$$

Які з оцінок та яких параметрів є незміщеними оцінками?

$$\hat{\theta}_1 = \min\{\zeta_1, \zeta_2, \ldots, \zeta_n\};$$

$$\hat{\theta}_2 = \max\{\zeta_1, \zeta_2, \ldots, \zeta_n\};$$

$$\hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n}\zeta_i;$$

$$\hat{\theta}_n = (\zeta_{n-1} + \zeta_n)/2.$$

Розв`язання. *Спочатку знайдемо розподіл випадкової величини* $\hat{\theta}_2 = \max\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$. *Оскільки всі випадкові величини* $\zeta_1, \zeta_2, \ldots, \zeta_n$ *незалежні, то, відповідно до теореми множення ймовірностей*

$$F_{\hat{\theta}_2}(x) = P\left(\hat{\theta}_2 < x\right) = P\left\{\max\{\zeta_1, \zeta_2, \ldots, \zeta_n\} < x\right\} =$$

$$= P\{\zeta_1 < x \cap \zeta_2 < x \cap \ldots \cap \zeta_n < x\} = \prod_{i=1}^{n} P\{\zeta_1 < x\}.$$

Тепер врахуємо, що всі випадкові величини $\zeta_1, \zeta_2, \ldots, \zeta_n$ *однаково розподілені, тобто*

$$\mathrm{P}\{\zeta_1 < x\} = \mathrm{P}\{\zeta_2 < x\} = \ldots = \mathrm{P}\{\zeta_n < x\}$$

або

$$F_{\zeta_1}(x) = F_{\zeta_2}(x) = \ldots = F_{\zeta_n}(x).$$

Тоді

$$F_{\hat{\theta}_2}(x) = \prod_{i=1}^{n} F_{\zeta_i}(x) = \left[F_{\zeta}(x)\right]^n = \left[\int_{-\infty}^{x} f(y, a, b)\,dy\right]^n.$$

За відомою функцією розподілу випадкової величини $\hat{\theta}_2$ *знайдемо густину її розподілу. За означенням*

$$f_{\hat{\theta}_2}(x) = \frac{d}{dx}F_{\hat{\theta}_2}(x) = nf(x, a, b)\left[\int_{-\infty}^{x} f(y, a, b)\,\mathrm{d}\,y\right]^{n-1},$$

оскільки

$$\int_{-\infty}^{x} f\left(y,a,b\right)dy=\begin{cases} 0,\ if\ \ x<a, \\ \dfrac{x-a}{b-a},\ if\ \ \mathrm{a}\le \mathrm{x}\le b, \\ 1, if\ \ x>b. \end{cases}$$

Отже,

$$f_{\hat{\theta}_2}\left(x\right)=\begin{cases} \dfrac{n}{\left(b-a\right)^n}\left(x-a\right)^{n-1}, if\ \ x\in\left[a,b\right], \\ 0,\ if\ \ \mathrm{x}\notin\left[a,b\right]. \end{cases}$$

За відомою щільністю розподілу оцінки $\hat{\theta}_2$ *можемо обчислити її математичне сподівання:*

$$M\hat{\theta}_2=\int_{-\infty}^{\infty} xf_{\hat{\theta}_2}\left(x\right)dx=\int_{a}^{b} x\frac{n}{\left(b-a\right)^n}\left(x-a\right)^{n-1}dx=\frac{n}{n+1}b+\frac{a}{n+1}\ .$$

Таким чином, оцінка $\hat{\theta}_2$ *не є незміщеною оцінкою ні параметра* a *, ні параметра* b*, проте вона є асимптотично незміщеною оцінкою параметра* b*. Дійсно*

$$\lim_{n\to\infty}\mathrm{M}\,\hat{\theta}_2=\lim_{n\to\infty}\left[\frac{n}{n+1}b+\frac{a}{n+1}\right]=b\,.$$

Знайдемо тепер розподіл випадкової величини $\hat{\theta}_1=\min\left\{\zeta_1,\zeta_2,\ldots,\zeta_n\right\}$*, тобто функцію*

$$F_{\hat{\theta}_1}\left(x\right)=P\left\{\hat{\theta}_1<x\right\}=P\left\{\min\left\{\zeta_1,\zeta_2,\ldots,\zeta_n\right\}<x\right\}.$$

Якщо за аналогією з попереднім розглядом перейти далі від розподілу ймовірностей для випадкової величини $\hat{\theta}_1$ *до функцій* *розподілу випадкових величин* $\zeta_1, \zeta_2, \ldots, \zeta_n$, *то виникає незручність, пов'язана з тим, що з умови* $\min\{\zeta_1, \zeta_2, \ldots, \zeta_n\} < x$ *не випливають умови* $\zeta_1, \zeta_2, \ldots, \zeta_n < x$. *Оскільки*

$$\mathrm{P}\left\{\hat{\theta}_1 < x\right\} + \mathrm{P}\left\{\hat{\theta} \geq x\right\} = 1,$$

то

$$F_{\hat{\theta}_1}(x) = 1 - P\left\{\hat{\theta}_1 \geq x\right\} = 1 - \prod_{i=1}^{n} P\{\zeta_i \geq x\} = 1 - \left[\int_x^{\infty} f(y,a,b)\,dy\right]^n .$$

Обчислимо інтеграл

$$\int_x^{\infty} f(y,a,b)\,dy = \begin{cases} 1, if\ x < a, \\ \dfrac{b-x}{b-a}, if\ a \leq x \leq b, \\ 0, if\ x > b. \end{cases}$$

Тепер

$$f_{\hat{\theta}_1}(x) = \frac{d}{dx} F_{\hat{\theta}_1}(x) = nf(x,a,b)\left[\int_x^{\infty} f(y,a,b)\,dy\right]^{n-1} =$$

$$= \begin{cases} \dfrac{n}{(b-a)^n}, if\ x \in [a,b], \\ 0, if\ \mathrm{x} \notin [a,b]. \end{cases}$$

За відомою щільністю розподілу оцінки $\hat{\theta}_1$ *обчислимо її математичне сподівання*

$$M\hat{\theta}_1 = \int_{-\infty}^{\infty} x f_{\hat{\theta}_1}(x)\,dx = \int_a^b x \frac{n}{(b-a)^n}(b-x)^{n-1}\,dx = \frac{n}{n+1}a + \frac{b}{n+1},$$

тобто $\hat{\theta}_1$ *не є незміщеною оцінкою ні параметра* a, *ні параметра* b, *але вона є асимптотично незміщеною оцінкою параметра* a. *Дійсно*

$$\lim_{n\to\infty} M\hat{\theta}_1 = \lim_{n\to\infty}\left[\frac{n}{n+1}a + \frac{b}{n+1}\right] = a.$$

Обчислимо математичне сподівання оцінки $\hat{\theta}_3$*. Оскільки випадкова величина* $\hat{\theta}_3$ *є сумою випадкових величин* $\zeta_1, \zeta_2, \ldots, \zeta_n$*, функції розподілу яких відомі, так само як і їх математичних очікувань*

$$M\zeta_1 = \frac{a+b}{2},$$

то скориставшись властивістю математичного очікування, а саме тим, що математичне сподівання суми дорівнює сумі математичних сподівань окремих доданків, маємо

$$M\hat{\theta}_3 = M\frac{1}{n}\sum_{i=1}^{n}\zeta_i = \frac{1}{n}\sum_{i=1}^{n}M\zeta_i = \frac{a+b}{2}.$$

Отже, $\hat{\theta}_3$ *є незміщеною оцінкою математичного сподівання рівномірного на відрізку* $[a,b]$ *розподілу. Аналогічно попередньому*

$$M\hat{\theta}_n = \frac{M\left(\zeta_{n-1} + \zeta_n\right)}{2} = \frac{a+b}{2},$$

тобто оцінка θ_n *також є незміщеною оцінкою математичного сподівання рівномірного на відрізку* $[a,b]$ *розподілу.*

Корисність теоретичної оцінки незміщеності оцінок $\hat{\theta}_1$*,* $\hat{\theta}_2$*,* $\hat{\theta}_3$ *і* $\hat{\theta}_n$ *очевидна. Асимптотична незміщеність оцінок* $\hat{\theta}_1$ *і* $\hat{\theta}_2$ *свідчить про доцільність для підвищення точності оцінок параметрів* a *і* b *збільшення обсягу вибірки* n*, оскільки зміщення оцінки щодо точного значення відповідного параметра прямує при цьому до нуля. Розглянемо тепер оцінки* $\hat{\theta}_3$ *і* $\hat{\theta}_n$*. Обидві ці оцінки є незміщеними оцінками математичного сподівання розподілу. Разом з тим, інтуїтивно очевидною є* *впевненість, що оцінка $\hat{\theta}_3$ є кращою за оцінку $\hat{\theta}_n$. математичним підґрунтям для такої впевненості, основаної на інтуїції, є така властивість оцінки як ефективність. Отже оцінимо ефективність оцінок $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$ і $\hat{\theta}_n$.*

Почнемо з перевірки того, чи буде оцінка $\hat{\theta}_2$ збігатися асимптотично за ймовірністю до b. Маємо

$$P\left\{\left|\hat{\theta}_2 - b\right| > \varepsilon\right\} = P\left\{\hat{\theta}_2 \in \left(-\infty, b-\varepsilon\right) \cup \left(b+\varepsilon, \infty\right)\right\} =$$

$$= \int_{-\infty}^{b-\varepsilon} f_{\hat{\theta}_2}(x)\,dx + \int_{b+\varepsilon}^{\infty} f_{\hat{\theta}_2}(x)\,dx = \int_{a}^{b-\varepsilon} f_{\hat{\theta}_2}(x)\,dx =$$

$$= \int_{a}^{b-\varepsilon} \frac{n}{(b-a)^n}(x-a)^{n-1}\,dx = \left(1 - \frac{\varepsilon}{b-a}\right)^n .$$

Звідси

$$\lim_{n\to\infty} P\left\{\left|\hat{\theta}_2 - b\right| > \varepsilon\right\} = \lim_{n\to\infty}\left(1 - \frac{\varepsilon}{b-a}\right)^n = 0$$

і оцінка $\hat{\theta}_2$ є слушною оцінкою параметра θ. Так само слушною є і оцінка θ. Слушність оцінки $\hat{\theta}_3$ випливає з факту існування закону великих чисел, сформульованого саме для випадкових величин, що мають структуру, подібну до $\hat{\theta}_3$. Навпаки, оцінка $\hat{\theta}_n$ є неслушною.

Метод найбільшої правдоподібності

Метод найбільшої правдоподібності використовується для знаходження оцінок параметрів вибірки. В основі цього методу лежить проста ідея, а саме, що спостережуваним елементам вибірки відповідає найбільша ймовірність появи при вимірюваннях саме таких їх значень. Тобто, якщо розглянути вибірку $\mathbf{x}=(x_1,x_2,...,x_n)$, яка є однією з можливих реалізацій випадкового вектора $\boldsymbol{\xi}=\left(\xi_1,\xi_2,...,\xi_n\right)$, де кожній компоненті вектора відповідає одна і та сама густина ймовірності $P_\xi(x,\theta)$, то густиною ймовірності для випадкового вектора $\boldsymbol{\xi}=\left(\xi_1,\xi_2,...,\xi_n\right)$ буде функція $P_{\boldsymbol{\xi}}(\mathbf{x},\theta)=P_{\xi_1}(x_1,\theta)P_{\xi_2}(x_2,\theta)...P_{\xi_n}(x_n,\theta)$. Максимальне значення ця густина ймовірності матиме для $\xi_1=x_1$, $\xi_2=x_2$, ..., $\xi_n=x_n$. Таке максимальне значення називається функцією правдоподібності

$$L(\mathbf{x},\theta)=P_{\xi_1}(x_1,\theta)P_{\xi_2}(x_2,\theta)...P_{\xi_n}(x_n,\theta).$$

Очевидно, що функція правдоподібності є функцією невідомого параметру θ. В якості найкращої оцінки цього параметру в методі найбільшої правдоподібності беруть екстремум (максимум) функції правдоподібності вибірки. Замість екстремуму функції правдоподібності зручно шукати екстремум її логарифму, оскільки структура останньої функції простіша, а критичні точки ті самі. Отже рівняння правдоподібності збігається з необхідною умовою екстремуму функції і має вигляд

$$\frac{dL(\mathbf{x},\theta)}{d\theta}=0.$$

Розв'язки цього рівняння називаються оцінками найбільшої правдоподібності для параметру θ.

Якщо невідомих параметрів декілька. Нехай це $\boldsymbol{\theta}=(\theta_1,\theta_2,...,\theta_r)$, то замість одного рівняння правдоподібності, слід розв'язати систему таких рівнянь

$$\frac{\partial L(\mathbf{x},\boldsymbol{\theta})}{\partial\theta_k}=0,\, k=1,2,...,r\,/$$

Задача 1. *Нехай величини* $x_1, x_2, ..., x_k$ *мають нормальний розподіл, а невідомими параметрами є їх математичне очікування* $a = Mx_k$ *і дисперсія* $b = \sigma^2 = \mathrm{D}\, x_k$. *Знайти їх оцінку найбільшої правдоподібності.*

Розв`язання. *Функція найбільшої правдоподібності має вигляд*

$$L = L\left(x_1, ..., x_n, a, b\right) = \left(\frac{1}{\sqrt{2\pi b}}\right)^n \exp\left(-\sum_{k=1}^{n} \frac{\left(x_k - a\right)^2}{2b}\right).$$

Відповідно

$$\ln(L) = -\frac{n}{2}\left(\ln 2\pi + \ln b\right) - \frac{1}{2b}\sum_{k=1}^{n}\left(x_k - a\right)^2 .$$

Для оцінок $\tilde{a}$ *і* $\tilde{b}$ *параметрів* a, b *отримаємо наступну систему рівнянь правдоподібності*

$$\frac{\partial \ln(L)}{\partial a} = \frac{1}{\tilde{b}}\sum_{k=1}^{n}\left(x_k - \tilde{a}\right) = 0,$$

$$\frac{\partial \ln(L)}{\partial b} = -\frac{n}{2\tilde{b}} + \frac{1}{2\tilde{b}^2}\sum_{k=1}^{n}\left(x_k - \tilde{a}\right)^2 = 0.$$

вибірки. Перше рівняння можна записати так

$$\sum_{k=1}^{n} x_k - n\tilde{a} = 0.$$

Звідси

$$\tilde{a} = \bar{a} = \frac{1}{n}\sum_{k=1}^{n} x_k .$$

Для другого рівняння маємо

$$-n + \frac{1}{\tilde{b}}\sum_{k=1}^{n}\left(x_k - \tilde{a}\right)^2 = 0$$

Звідки

$$\tilde{b} = \bar{b} = s^2 = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \tilde{a}\right)^2 .$$

Отже, вибіркове середнє є найкращою в сенсі методу найбільшої правдоподібності оцінкою математичного очікування, а вибіркова дисперсію є аналогічною оцінкою дисперсії.

Задача 2. *Нехай величини* $x_1, x_2, ..., x_k$ *мають закон розподілу Бернуллі, а невідомим параметром є ймовірність успіху* p. *Знайти її оцінку найбільшої правдоподібності.*

Розв`язання. *Оскільки ймовірність* m *успіхів при* n *випробуваннях в схемі Бернуллі визначається законом розподілу*

$$P_n(m) = C_n^m p^m (1-p)^{n-m} ,$$

то функція найбільшої правдоподібності має вигляд

$$L = L(m_1, m_2, ..., m_n) = (C_n^{m_1} C_n^{m_2} ... C_n^{m_n})\, p^{m_1+m_2+...,+m_n} (1-p)^{n^2-(m_1+m_2+...,+m_n)} .$$

Відповідно

$$\ln(L) = \ln\left[(C_n^{m_1} C_n^{m_2} ... C_n^{m_n})\right] + (m_1 + m_2 + ..., +m_n)\ln(p) +$$

$$+\left[n^2 - (m_1 + m_2 + ..., +m_n)\right]\ln(1-p) .$$

Для оцінки $\hat{p}$ *параметру* p *отримаємо наступне рівняння правдоподібності вибірки*

$$\frac{\partial \ln(L)}{\partial p} = \frac{m_1 + ... + m_n}{\hat{p}} - \frac{n^2 - (m_1 + ... + m_n)}{1-\hat{p}} = 0$$

або

$$\frac{\bar{m}}{\hat{p}} - \frac{n-\bar{m}}{1-\hat{p}} = 0$$

Це рівняння має очевидний розв'язок

$$\hat{p} = \frac{\bar{m}}{n} ,$$

де

$$\overline{m} = \frac{m_1 + ... + m_n}{n}$$

- вибіркове середнє кількості успіхів.

Метод моментів

Можливо, найпростішим методом знаходження невідомих параметрів розподілів випадкових величин є метод моментів. Він полягає у порівнянні експериментальних моментів, знайдених на основі вибірки, і моментів теоретичних. Моменти, знайдені на основі вибірки, для кожної конкретної вибірки є числами. Теоретичні моменти є функціями невідомих параметрів розподілу ймовірності. Прирівнюючи відповідні моменти між собою ми отримуємо рівняння для знаходження невідомих параметрів розподілу ймовірностей. Розглянемо цей алгоритм детальніше. Нехай функція розподілу ймовірності $F_{\xi}(x,\lambda)$ випадкової величини ξ залежить від невідомого параметру λ. Нагадаємо, що вибіркові та теоретичні моменти визначаються так:

$$a_{\nu} = \frac{1}{n}\sum_{i=1}^{n} x_i^{\nu}$$

- вибіркові початкові моменти для дискретної випадкової величини,

$$m_{\nu} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^{\nu}$$

- центральні вибіркові моменти для дискретної випадкової величини,

$$\alpha_{\nu}(\lambda) = M\xi^{\nu} = \int_{-\infty}^{\infty} x^{\nu} P_{\xi}(x,\lambda)dx \,,$$

$$\alpha_{\nu}(\lambda) = M\xi^{\nu} = \int x^{\nu} dF_{\xi}(x,\lambda)$$

- теоретичні початкові моменти для неперервної випадкової величини та

$$\mu_\nu(\lambda) = M(\xi - M\xi)^\nu = \int_{-\infty}^{\infty} (x - M\xi)^\nu P_\xi(x,\lambda)dx\,,$$

$$\mu_\nu(\lambda) = M(\xi - M\xi)^\nu = \int (x - M\xi)^\nu dF_\xi(x,\lambda)$$

- центральні теоретичні моменти для неперервної випадкової величини.

Якщо функція розподілу ймовірностей залежить лише від одного невідомого параметру λ, то досить прирівняти перший теоретичний і вибірковий моменти для знаходження цього параметру

$$a_1 = \alpha_1(\lambda)$$

або

$$m_1 = \mu_1(\lambda).$$

Якщо таких параметрів декілька: $\lambda_1, \lambda_2, ..., \lambda_k$, то слід прирівняти відповідну кількість перших моментів

$$a_1 = \alpha_1(\lambda_1, \lambda_2, ..., \lambda_k)\,,$$

$$a_2 = \alpha_2(\lambda_1, \lambda_2, ..., \lambda_k)\,,$$

$$......................$$

$$a_k = \alpha_k(\lambda_1, \lambda_2, ..., \lambda_k)$$

або

$$m_1 = \mu_1(\lambda_1, \lambda_2, ..., \lambda_k)\,,$$

$$m_2 = \mu_2(\lambda_1, \lambda_2, ..., \lambda_k)\,,$$

$$......................$$

$$m_k = \mu_k(\lambda_1, \lambda_2, ..., \lambda_k)$$

або частину перших і частину других. Тривіальним наслідком методу моментів є такий результат для нормального розподілу, що залежить від двох параметрів

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

В якості рівнянь для визначення цих параметрів можна взяти наступні два рівняння

$$a_1=\alpha_1(a,\sigma)\,,\qquad m_2=\mu_2(a,\sigma)\,.$$

Оскільки

$$a_1=\frac{1}{n}\sum_{i=1}^{n}x_i\,,$$

$$m_2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2\,,$$

а

$$\alpha_1(a,\sigma)=a\,,\ \ \mu_2(a,\sigma)=\sigma^2\,,$$

то отримуємо популярні для математичного очікування і дисперсії нормального розподілу оцінки у вигляді вибіркового середнього і вибіркової дисперсії відповідно

$$a=\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i\,,$$

$$\sigma^2=s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2\,.$$

3.8.Інтервальні оцінки

Точкові оцінки дають наближене значення невідомого параметра. Але вони нічого не говорять про точність оцінок. Оскільки будь-яка оцінка є випадковою величиною, то вона може приймати практично довільні значення. Проте, в загальному випадку, різні значення оцінка може набувати з різною ймовірністю. Для деякого значення оцінки ймовірність може бути максимальною. Для інших значень вона спадатиме в міру відхилення від такого екстремального значення. Зручно накривати таке екстремальне значення випадкової величини певним інтервалом і оцінювати ймовірність, з якою можливі значення випадкової величини належать цьому інтервалу. Такий опис точності оцінки називається інтервальною оцінкою. Тобто математична статистика пропонує такий спосіб визначення точності оцінки: зазначення інтервалу, якому належить оцінка випадкової величини з одночасним зазначенням ймовірності, з якою цей факт має місце. Насправді це найкращий з можливих способів оцінки похибки оцінки.

Певним компромісом є оцінка похибки у термінах абсолютної похибки. У цьому разі в якості абсолютної похибки можна узяти середньоквадратичне відхилення (корінь квадратний з вибіркової дисперсії) вибірки, з допомогою якої ми оцінюємо невідомий параметр функції розподілу ймовірності. Тоді ми просто стверджуємо, що з переважаючою ймовірністю, якою саме ми не знаємо, можливі значення оцінки належать інтервалу, довжина якого є подвоєним середнім квадратичним відхиленням значень оцінки від її середнього значення. У тому ж разі, якщо функція розподілу елементів вибірки відома, можна виконати найкращу з можливих оцінок похибки - інтервальну оцінку, про яку говорилось вище.

Перейдемо до точних означень. Розглянемо вибірку $x_1, x_2, ..., x_n$ з функцією розподілу $P\left(x_i < x\right) = F_\xi\left(x, \theta\right)$, де θ – невідомий параметр. Інтервал $(\underline{\theta}, \overline{\theta})$, у якому з імовірністю $1-\alpha$ лежить невідоме значення параметру θ називається довірчим інтервалом.

Розглянемо декілька випадків, коли довірчий інтервал можна легко знайти. Ці випадки стосуються нормального розподілу для елементів вибірки, а також деяких простих функцій елементів вибірки, функції яких можна знайти виходячи з нормального розподілу.

Довірчий інтервал для математичного очікування при відомій дисперсії

Нехай елементи вибірки розподілені нормально з параметрами $\left(a,\sigma^2\right)$

$$P_\xi(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-a)^2}{2\sigma^2}\right].$$

При цьому параметр a невідомий, а параметр σ^2 відомий. Знайдемо довірчий інтервал для a. Випадкова величина (вибіркове середнє) $\overline{x}=\left(x_1+...+x_n\right)/n$, яка є однією з можливих реалізацій випадкового вектора $\xi=(\xi_1,\xi_2,...,\xi_n)$, у зазначеному випадку теж має нормальний розподіл з параметрами $\left(a,\sigma^2/n\right)$

$$P_{\overline{x}}(x)=\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(\overline{x}-a)^2}{2\sigma^2/n}\right],$$

тобто $M\overline{x}=a$, $D\overline{x}=\sigma^2/n$. Це точний результат за будь-якого обсягу вибірки і наближений у разі іншого закону розподілу випадкової величини, якщо вибірка достатньо велика. Наведена густина розподілу є доволі складною функцією густини ймовірності для простої за конструкцією випадкової величини. Вона залежить як від елементів вибірки, так і від невідомого параметра розподілу. Цю функцію розподілу можна суттєво спростити, звівши її до стандартного нормального розподілу з нульовим математичним очікуванням і одиничною дисперсією. Але, при цьому, це вже буде функція густини розподілу ймовірностей для значно складнішої випадкової величини. Виконаємо кілька простих перетворень

$$P_{\bar{x}-a}(x)=\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\exp\left[-\frac{x^2}{2\sigma^2/n}\right],$$

$$P_{\frac{\bar{x}-a}{\sigma/\sqrt{n}}}(x)=\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x^2}{2}\right].$$

Таким чином для випадкової величини

$$\eta=\frac{\bar{x}-a}{\sigma/\sqrt{n}},$$

Ми отримуємо стандартний нормальний закон розподілу

$$P_\eta(x)=\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x^2}{2}\right].$$

Від густини розподілу ймовірностей перейдемо до інтервальною оцінки випадкової величини. Очевидно, що ймовірність (довірча ймовірність) того, що значення випадкової величини η належать інтервалу (довірчому інтервалу) $[-\Delta,\Delta]$, буде такою

$$P\left(\eta<\Delta\right)=\frac{1}{\sqrt{2\pi}}\int_{-\Delta}^{\Delta}\exp(-x^2/2)dx$$

або

$$P\left(\eta<\Delta\right)=2\Phi_0(\Delta),$$

де

$$\Phi_0\left(\Delta\right)=\frac{1}{\sqrt{2\pi}}\int_{0}^{\Delta}\exp(-x^2/2)dx$$

Отже, єдиною незалежною змінною, від якої залежить довірча ймовірність є півширина нерозмірного довірчого інтервалу

$$P\left(\eta<\Delta\right)=2\Phi_0(\Delta).$$

Від довірчого інтервалу щодо η можна перейти до довірчого інтервалу щодо шуканого математичного очікування. Дійсно, нехай

$$\eta < \Delta = \left| \frac{\overline{x} - a}{\sigma / \sqrt{n}} \right| < \Delta$$

Цю нерівність для значень випадкової величини під знаком модуля можна записати як подвійну нерівність для значень самої випадкової величини

$$-\Delta < \frac{a - \overline{x}}{\sigma / \sqrt{n}} < \Delta .$$

Перестановка додатків у чисельнику не змінить цієї нерівності

$$-\Delta < \frac{a - \overline{x}}{\sigma / \sqrt{n}} < \Delta .$$

Нерівність не зміниться і при множенні всіх її частин на один і той самий додатний множник

$$-\Delta\sigma / \sqrt{n} < a - \overline{x} < \Delta\sigma / \sqrt{n}$$

Не зміниться нерівність і при додаванні до всіх її частин однієї і тої самої величини

$$\overline{x} - \Delta\sigma / \sqrt{n} < a < \overline{x} + \Delta\sigma / \sqrt{n} .$$

Отже, ми отримали довірчий інтервал щодо шуканого математичного очікування. Таким чином, задавшись наперед якимось значенням довірчої ймовірності, ми можемо отримати відповідний нерозмірний довірчий інтервал $[-\Delta, \Delta]$ для випадкової величини η. За знайденим значення Δ знаходимо розмірний довірчий інтервал для величини a.

І нерозмірний довірчий інтервал і довірча ймовірність не залежать ні від елементів вибірки, ні від її обсягу. У свою чергу розмірний довірчий інтервал залежатиме як від елементів вибірки, так і від її обсягу.

Довірча ймовірність звичайно мало відрізняється від одиниці і її частот записують у вигляді $1-2\alpha$, де 2α є малою величиною. Двійка у цьому виразі звичайно пишеться для того, щоб величину α зручніше було виразити через функцію Лапласа. Для випадку нормального розподілу легко записати зв'язок між α і Δ

$$\alpha = \frac{1}{\sqrt{2\pi}}\int_{\Delta}^{\infty} \exp(-x^2/2)dx$$

або

$$\alpha = 1/2 - \Phi_0(\Delta).$$

α можна інтерпретувати, як ймовірність приналежності значення випадкової величини інтервалу $[\Delta,\infty)$ або $(-\infty,-\Delta]$.

Якщо $\Delta = 1$, то довірча ймовірність

$$P\left\{\bar{x} - \frac{\sigma}{\sqrt{n}} < a < \bar{x} + \frac{\sigma}{\sqrt{n}}\right\} = 0.6826.$$

Якщо $\Delta = 2$, то довірча ймовірність

$$P\left\{\bar{x} - 2\frac{\sigma}{\sqrt{n}} < a < \bar{x} + 2\frac{\sigma}{\sqrt{n}}\right\} = 0.9548.$$

Якщо $\Delta = 3$, то довірча ймовірність

$$P\left\{\bar{x} - 3\frac{\sigma}{\sqrt{n}} < a < \bar{x} + 3\frac{\sigma}{\sqrt{n}}\right\} = 0.9972.$$

Майже достовірним є той факт, що випадкова величина не відхилиться від математичного сподівання (вибіркового середнього) більше, ніж на $3\sigma/\sqrt{n}$. Ширина довірчого інтервалу швидко зростає із зростанням довірчої ймовірності, але при нерозмірній ширині 3 довірча ймовірність вже настільки мало відрізняється від одиниці, що приналежність вимірюваної величини такому інтервалу можна вважати практично достовірною подією.

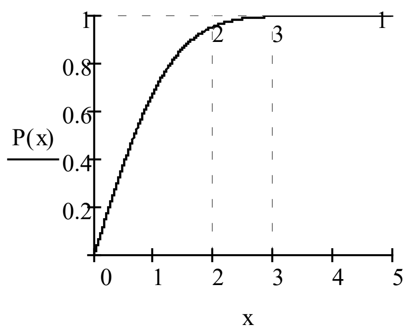
Співвідношення між нерозмірним довірчим інтервалом і довірчою ймовірністю вимагає певного компромісу між ними. Очевидно, що довірчій ймовірності одиниця відповідає лише нескінчений довірчий інтервал. Нульовому довірчому інтервалу відповідає і нульова довірча ймовірність. Нам потрібний якомога вужчий довірчий інтервал при доволі високій довірчій ймовірності. Значення $\Delta = 3$ є одним з тих, що якраз і відповідає цим вимогам.

Приклад 1. *Розглянемо алгоритм знаходження довірчого інтервалу і довірчої ймовірності на конкретному прикладі. Використаємо універсальне співвідношення між довірчим інтервалом і довірчою ймовірністю*

$$P(\Delta) := 2\Phi_0(\Delta),$$

$$\Phi_0(\Delta) := \frac{1}{\sqrt{2\pi}} \int_0^{\Delta} \exp\left(-\frac{x^2}{2}\right) dx$$

Графік останньої функції наступний



Мал. 1. *Залежність довірчої ймовірності від півширини нерозмірного довірчого інтервалу.*

Сформуємо тепер вибірку достатньо великого обсягу, наприклад, $n := 100$*, при відомій дисперсії , наприклад, 4,* $\sigma := 2$. *Для цього використаємо прикладний пакет Mathcad і вбудований в нього генератор випадкових нормально розподілених чисел*

$$\xi := rnorm(100, 10, 2)$$

$\xi^T =$

	0	1	2	3	4	5	6	7
0	9.122	8.641	9.053	8.097	6.629	10.087	9.759	...

Пакет виводить випадкові числа у вигляді стовпця. Для економії місця ми транспонували цей стовпець у рядок.

Тепер ми можемо записати розмірний довірчий інтервал саме для даної вибірки, що відповідає результатам вимірювання конкретної фізичної величини Для цього нам потрібне вибіркове середнє

$$\overline{\xi}(n) := \frac{1}{n}\sum_{k=0}^{n-1}\xi_k\,, \qquad \overline{\xi}(100) = 9.873\,.$$

Тепер для лівого і правого кінців інтервалу матимемо

$$\overline{\xi}(n) - \Delta\frac{\sigma}{\sqrt{n}}, \quad \overline{\xi}(n) + \Delta\frac{\sigma}{\sqrt{n}}.$$

Якщо надати нерозмірній півширині інтервалу значень: $\Delta := 1$, $\Delta := 2$, $\Delta := 3$, *то відповідні розмірні довірчі інтервали будуть:* $[9.673, 10.073]$, $[9.473, 10.273]$, $[9.273, 10.473]$, *а довірчі ймовірності* $P(1) = 0.6826$, $P(2) = 0.9548$, $P(3) = 0.9972$. *Насправді, нам відоме точне значення математичного очікування. Коли ми створювали вибірку за допомогою генератора випадкових чисел, то в якості одного з аргументів ми задали це математичне очікування* $M\xi := 10$. *Видно, що за даним обсягом вибірки точне значення математичного очікування не потрапило у довірчий інтервал. Тобто обсяг вибірки виявився недостатнім навіть для нерозмірної півширини інтервалу* $\Delta := 3$.

Далі звузити розмірні довірчі інтервали при незмінній довірчих ймовірностях можна за рахунок збільшення обсягу вибірки. Нехай $n := 1000$. *Тоді*

$$\xi := rnorm(1000, 10, 2),$$

$\xi^T =$

	0	1	2	3	4	5	6	7
0	10.345	6.641	7.887	11.012	13.493	12.555	13.112	...

$$\xi\xi(1000) = 10.095$$

Якщо надати нерозмірній півширині інтервалу значень: $\Delta := 1$, $\Delta := 2$, $\Delta := 3$, *то відповідні розмірні довірчі інтервали* *будуть:* $[9.895, 10.295]$, $[9.695, 10.495]$, $[9.495, 10.695]$, *а довірчі ймовірності* $P(1) = 0.6826$, $P(2) = 0.9548$, *удесятеро, ми зменшили ширини розмірних довірчих інтервалів приблизно у* $\sqrt{10}$ *разів при незмінних довірчих ймовірностях.*

Висновок з даного розгляду простий: обсяг вибірки має бути дуже великим для реалізації високої довірчої ймовірності і вузького довірчого інтервалу.

Легко можна переконатись, що використавши вибірки для випадкової величини з іншим законом розподілу, що має скінчену дисперсію, з тим самим математичним очікуванням, наприклад, рівномірним, при великому обсязі вибірки ми отримає для тих самих довірчих ймовірностей дуже близькі розмірні довірчі інтервали. При цьому ми використовуватимемо попередні формули для нормально розподіленої випадкової величини. Нехай

$P(3) = 0.9972$. *Для даного обсягу вибірки точне значення математичного очікування, яке нам було відоме заздалегідь, потрапило у всі три розмірні довірчі інтервали. Збільшивши обсяг вибірки*

$$\xi := runif(1000, 5, 15),$$

$\xi^T =$

	0	1	2	3	4	5	6	7
0	14.583	14.49	13.351	6.152	11.28	8.084	8.2	...

Якщо надати нерозмірній півширині інтервалу значень: $\Delta := 1$, $\Delta := 2$, $\Delta := 3$, *то відповідні розмірні довірчі інтервали будуть:* $[9.828, 10.228]$, $[9.628, 10.428]$, $[9.428, 10.628]$, *а довірчі ймовірності* $P(1) = 0.6826$, $P(2) = 0.9548$, $P(3) = 0.9972$. *Отримані довірчі інтервали лише у третій значущій цифрі відрізняються від попередніх результатів, отриманих для того самого обсягу вибірки, але для нормального розподілу.*

Довірчий інтервал для математичного очікування при невідомій дисперсії

Складнішим, але і реалістичнішим є випадок, коли обидва параметри a і σ^2 є невідомими. У цьому разі з двох випадкових величин $\overline{x}$ і S^2 (вибіркового середнього і виправленої вибіркової дисперсії) можна утворити третю випадкову величину

$$\eta = \frac{\overline{x} - a}{S / \sqrt{n-1}},$$

де і у чисельнику, і у знаменнику присутні випадкові величини. Порівняємо з випадком відомої дисперсії, коли знаменник є числом

$$\eta = \frac{\overline{x} - a}{\sigma / \sqrt{n}}.$$

Нагадаємо, що $M\overline{x} = M\xi$, $MS^2 = D\xi / n$. Як відомо, для елементів вибірки розподілених за нормальним законом, випадкова величина $\eta = (\overline{x} - a)/(S/\sqrt{n-1})$ розподілена за законом Стьюдента з $n-1$ степенями свободи. Це твердження є точним результатом. У разі, якщо випадкова величина розподілена за іншим законом, але її дисперсія скінчена, то це твердження буде тим точнішим, чим більшим буде обсяг вибірки. Ми говоримо про $n-1$ ступінь вільності через те, що для вибірки обсягом n використовуємо не вибіркову дисперсію, а виправлену вибіркову дисперсію. Також, як і у разі нормального розподілу, для введеної випадкової величини, що описується розподілом Стьюдента, ні межі довірчого інтервалу, ні довірча ймовірність не залежать від шуканих параметрів. Натомість з'являється залежність довірчої ймовірності від обсягу вибірки n. У разі розподілу Стьюдента також легко записати зв’язок між довірчим інтервалом і довірчою ймовірністю.

Густина розподілу Стьюдента з $n-1$ ступенями вільності має вигляд

$$P_\eta(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)}\left(1+\frac{x^2}{n-1}\right)^{-\frac{n}{2}}, \quad x \in (-\infty, \infty).$$

У нашому випадку

$$\eta = \frac{\bar{x} - a}{S / \sqrt{n-1}}.$$

Її відмінність від нормального розподілу у тому, що вона суттєвим чином залежить від обсягу вибірки.

Інтервальна оцінка приналежності значень випадкової величини η інтервалу $[-\Delta, \Delta]$ визначиться подібно випадку нормального розподілу

$$P\left(\eta < \Delta\right) = \int_{-\Delta}^{\Delta} P_\eta(x)dx.$$

Очевидно, що довірча ймовірність тепер буде функцією як нерозмірної півширини довірчого інтервалу, так і обсягу вибірки

$$P\left(\eta < \Delta\right) = P(\Delta, n).$$

Розподіл Стьюдента не є парною функцією, як нормальний розподіл, тому від меж інтегрування $-\Delta, \Delta$ не можна перейти до меж інтегрування $0, \Delta$. Ми знову отримуємо універсальне співвідношення між довірчою ймовірністю і нерозмірним довірчим інтервалом, що не залежить від елементів вибірки, а лише від її обсягу. Наприклад, для довірчої ймовірності $P(\Delta, 11) = 0.99$, отримаємо $\Delta = 3.169$. При переході від нерозмірного довірчого інтервалу $-\Delta < \eta < \Delta$ до розмірного довірчого інтервалу для вимірюваного параметру матимемо

$$\bar{x} - \Delta\frac{S}{\sqrt{n-1}} < a < \bar{x} + \Delta\frac{S}{\sqrt{n-1}}.$$

Алгоритм реалізації інтервальною оцінки математичного очікування при відомій дисперсії аналогічний попередньому випадку. Ми визначаємось з тим, яке значення довірчої ймовірності нас влаштовує. З універсального співвідношення між довірчою ймовірністю і нерозмірною півшириною довірчого інтервалу ми визначаємо цю ширину. Пам’ятаємо, що ця ширина залежатиме від обсягу вибірки. Від нерозмірного довірчого інтервалу переходимо до розмірного довірчого інтервалу для вимірюваного математичного очікування за наведеною вище формулою.

Розподіл Стьюдента переходить в нормальний розподіл у межі $n \to \infty$. На практиці це означає, що для вибірок обсягом більше тридцяти можна використовувати нормальний розподіл ймовірності, для менших вибірок - розподіл Стьюдента.

Всі наведені оцінки точні лише у разі нормального розподілу елементів вибірки. Якщо ж розподіл ймовірності відрізняється від нормального, але має скінчену дисперсію, то все одно залишається доцільним використання розподілу Стьюдента для розрахунків довірчого інтервалу і довірчої ймовірності. У цьому разі всі оцінки будуть вже наближеними, але точність наближення швидко зростатиме при збільшенні обсягу вибірки. Надійним керманичем у цих міркуваннях безумовно є центральна межова теорема. Дійсно, відповідно до неї, незалежно від функції розподілу елементів вибірки, вибіркове середнє при великому обсягу вибірки описуватиметься нормальним законом розподілу, а вибіркова дисперсія - законом Стьюдента.

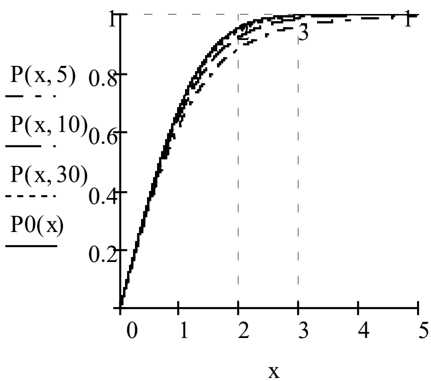
Приклад 2. *Розглянемо алгоритм знаходження довірчого інтервалу і довірчої ймовірності на конкретному прикладі. Використаємо універсальне співвідношення між довірчим інтервалом і довірчою ймовірністю*

$$P(\Delta,n)=\int_{-\Delta}^{\Delta}\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)}\left(1+\frac{x^2}{n-1}\right)^{-\frac{n}{2}}dx\,.$$

Для порівняння наведемо і відповідний зв'язок для нормального розподілу

$$P(\Delta,n)=\int_{-\Delta}^{\Delta}\frac{1}{\sqrt{2\pi}}\exp\left(-x^2/2\right)\,dx\,.$$

Графік останньої функції наступний для різних значень довірчого інтервалу у порівнянні з графіком нормального розподілу має вигляд



Мал. 1. *Залежність довірчої ймовірності від півширини нерозмірного довірчого інтервалу.*

З малюнка видно, що густина розподілу Стьюдента помітно залежить від обсягу вибірки для малих обсягів вибірки. Для обсягу вибірки 30 різниці між нормальним розподілом і розподілом Стьюдента візуально вже не видно.

Сформуємо тепер вибірку достатньо великого обсягу, наприклад, $n:=100$, *при відомій дисперсії , наприклад, 4,* $\sigma:=2$. *Для цього використаємо прикладний пакет Mathcad і вбудований в нього генератор випадкових нормально розподілених чисел*

$$\xi := rnorm(100,10,2)$$

$\xi^T =$

	0	1	2	3	4	5	6	7
0	9.122	8.641	9.053	8.097	6.629	10.087	9.759	...

Пакет виводить випадкові числа у вигляді стовпця. Для економії місця ми транспонували цей стовпець у рядок.

Тепер ми можемо записати розмірний довірчий інтервал саме для даної вибірки, що відповідає результатам вимірювання конкретної фізичної величини. Для цього нам потрібне вибіркове середнє

$$\overline{\xi}(n) := \frac{1}{n}\sum_{k=0}^{n-1}\xi_k\,, \quad \overline{\xi}(100) = 10.204\,.$$

Також нам потрібне виправлена вибіркова дисперсія

$$S2(n) := \frac{1}{n-1}\sum_{k=0}^{n-1}(\xi_k - \overline{\xi}(n))^2\,, \quad S2(100) = 4.203\,.$$

Тепер для лівого і правого кінців розмірного довірчого інтервалу матимемо

$$\overline{\xi}(n) - \Delta\frac{\sqrt{S2(n)}}{\sqrt{n-1}}, \quad \overline{\xi}(n) + \Delta\frac{\sqrt{S2(n)}}{\sqrt{n-1}}.$$

Якщо надати нерозмірній півширині інтервалу значень: $\Delta := 1$*,* $\Delta := 2$*,* $\Delta := 3$*, то відповідні розмірні довірчі інтервали будуть:* $[9.998, 10.41]$*,* $[9.473, 10.273]$*,* $[9.273, 10.473]$*, а довірчі ймовірності* $P(1) = 0.6826$*,* $P(2) = 0.9548$*,* $P(3) = 0.9972$*. Насправді, нам відоме точне значення математичного очікування. Коли ми створювали вибірку за допомогою генератора випадкових чисел, то в якості одного з аргументів ми задали це математичне очікування* $M\xi := 10$*. Видно, що за даним обсягом вибірки точне значення математичного очікування не потрапило у довірчий інтервал. Тобто обсяг вибірки виявився недостатнім навіть для нерозмірної півширини інтервалу* $\Delta := 3$ *.*

Далі звузити розмірні довірчі інтервали при незмінній довірчих ймовірностях можна за рахунок збільшення обсягу вибірки. Нехай $n := 1000$*. Тоді*

$$\xi := rnorm(1000,10,2)\,,$$

$\xi^{T} =$

	0	1	2	3	4	5	6	7
0	10.345	6.641	7.887	11.012	13.493	12.555	13.112	...

$$\overline{\xi}(1000) = 10.095$$

Якщо надати нерозмірній півширині інтервалу значень: $\Delta := 1$, $\Delta := 2$, $\Delta := 3$, *то відповідні розмірні довірчі інтервали будуть:* $[9.774, 10.198]$, $[9.562, 10.409]$, $[9.35, 10.621]$, *а довірчі ймовірності* $P(1) = 0.680$, $P(2) = 0.952$, $P(3) = 0.997$. *Для даного обсягу вибірки точне значення математичного очікування, яке нам було відоме заздалегідь, потрапило у всі три розмірні довірчі інтервали. Використання розподілу Стьюдента для вибірки даного обсягу практично повторює результати використання нормального розподілу.*

Довірчий інтервал для дисперсії при відомому математичному очікуванні

Розглянемо тепер ситуацію, коли для вибірки $x_1, x_2, ..., x_n$ математичне очікування, тобто параметр a, відоме, а дисперсія, тобто параметр σ^2, ні. При цьому, як і раніше, елементи вибірки розподілені за нормальним законом. Знайдемо довірчий інтервал для параметра σ^2. Для відомого a скористаємось тим, що величина

$$\eta = (n-1)S^2 / \sigma^2$$

розподілена за законом хі-квадрат з $n-1$ ступенем вільності. Тут виправлена вибіркова дисперсія

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 .$$

Ця випадкова величина, відміну від попередніх випадків, набуває лише додатних значень. Для нормально розподіленої випадкової величини цей результат є точним. Для довільно розподіленої випадкової величини він є наближеним і ступень його точності тим вищий, чим більшим є обсяг вибірки.

Для розподілу хі-квадрат, густина ймовірності якого

$$P_\eta(x,n)=\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}x^{\frac{n}{2}-1}\exp\left(-\frac{x}{2}\right),\ x>0,\ \xi=\xi_1^2+\xi_2^2+...+\xi_n^2,$$

також дуже легко записати зв'язок між нерозмірним довірчим інтервалом $[a,b]$ і довірчою ймовірністю. Дійсно,

$$P\left(a<\eta<b\right)=\int_a^b P_\eta(x,n)dx.$$

Тобто довірча ймовірність є функцією меж нерозмірного довірчого інтервалу і обсягу вибірки і не залежить від елементів самої вибірки

$$P\left(a<\eta<b\right)=P\left(a,b,n\right).$$

Від довірчого інтервалу для нерозмірної випадкової величини можна перейти до довірчого інтервалу для розмірної величини дисперсії і середньоквадратичного відхилення. Для цього виконаємо ряд тотожних перетворень для подвійної нерівності

$$a<\frac{(n-1)S^2}{\sigma^2}<b,$$

$$\frac{a}{(n-1)S^2}<\frac{1}{\sigma^2}<\frac{b}{(n-1)S^2},$$

$$\frac{(n-1)S^2}{a}>\sigma^2>\frac{(n-1)S^2}{b}.$$

Таким чином, інтервал $\left(\frac{(n-1)S^2}{b},\frac{(n-1)S^2}{a}\right)$ є розмірним довірчим інтервалом для дисперсії, а інтервал $\left(\frac{\sqrt{n-1}S}{\sqrt{b}},\frac{\sqrt{n-1}S}{\sqrt{a}}\right)$ є довірчим інтервалом для середньоквадратичного відхилення.

Алгоритм реалізації наведених формул такий. Спочатку ми задаємось прийнятним значенням довірчої ймовірності. Після цього підбираємо межі нерозмірного довірчого інтервалу відповідно до цього значення довірчої ймовірності. Далі від нерозмірного довірчого інтервалу переходимо до розмірного.

Так само, як і розподіл Стьюдента, розподіл хі-квадрат при великих обсягах вибірки переходить у нормальний розподіл.

Приклад 3. *Розглянемо алгоритм знаходження довірчого інтервалу і довірчої ймовірності на конкретному прикладі. Спочатку задамо прийнятний рівень довірчої ймовірності і обсягу вибірки, наприклад, 0.997 і 100. Використаємо універсальне співвідношення між довірчим інтервалом і довірчою ймовірністю*

$$P(a,b,n) := \int_a^b \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right) dx .$$

Емпіричним шляхом підберемо межі довірчого інтервалу так, щоб довірча ймовірність дорівнювала 0.997. *Після декількох спроб отримуємо, що*

$$P(65,145,100) = 0.997 .$$

Отже, нерозмірний довірчий інтервал для нерозмірної випадкової величини η *буде* $(65,145)$. *Для шуканої дисперсії* σ^2 *розмірний довірчий інтервал буде таким* $(3.037, 73.406)$.

Тут ми використали вибірку обсягом 100 з нормально розподіленої випадкової величини з математичним очікуванням 10 і середньоквадратичним відхиленням 2

$$\xi := rnorm(100,10,2)$$

$\xi^{T} =$

	0	1	2	3	4	5
0	10.875	13.521	8.705	11.102	13.093	...

Точність інтервальною оцінки можна дещо підвищити, якщо ймовірність неналежності значень вимірюваного параметру довірчому інтервалу порівну поділити між інтервалами $[0,a]$ і $[b,\infty)$

$$\frac{1-P\left(a<\eta<b\right)}{2}=\int_{0}^{a}P_{\eta}(x)dx\,,$$

$$\frac{1-P\left(a<\eta<b\right)}{2}=\int_{b}^{\infty}P_{\eta}(x)dx\,.$$

Приклад 4. *Нехай довірча ймовірність дорівнює 0.997. Тоді для вибірки обсягом 100 матимемо*

$$\int_{0}^{63.2}P_{\eta}(x)dx=0.0015\,.$$

Отже, $a:=63.2$. *З умови*

$$\int_{147.4}^{250}P_{\eta}(x)dx=0.0015$$

матимемо $b:=147.4$. *Це дещо точніше, ніж у Прикладі 1за тих самих умов.*

3.9.Статистична перевірка гіпотез

Вважатимемо гіпотезою будь-яке припущення щодо функції розподілу вимірюваної випадкової величини або параметрах, від яких вона залежить. Простою гіпотезою вважається припущення, що функція розподілу вимірюваної випадкової величини збігається з якоюсь відомою функцією розподілу. На основі вибірки ми можемо побудувати емпіричну функцію розподілу. Для з'ясування її близькості до якоїсь певної функції розподілу потрібно керуватись якимось кількісним критерієм. Таких критеріїв можна придумати багато Найпопулярнішим з них є критерій Пірсона. Він вводиться наступним чином.

Нехай у нас є випадкова величина ξ, при вимірюванні якої ми отримуємо наступну вибірку $x_1, x_2, \dots, x_n$. Множину значень випадкової величини розбиваємо на r множин

$S_1, S_2, ..., S_r$ без спільних точок. Для цього нам потрібно $r+1$ точок $a_1, a_2, ..., a_{r+1}$. При цьому правий кінець відповідного інтервалу виключається з відповідної множини, а лівий включається. В останньому інтервалі включаються обидва кінці. Нехай $p_i, i = 0,1,...,r$ теоретичні ймовірність того, що випадкова величина ξ належить множині S_i. Також нам потрібна теоретична ймовірність p_0 того, що значення випадкової величини набувають значень, менших за найменший елемент вибірки. Потрібна і теоретична ймовірність p_{r+1} того, що значення випадкової величини набувають значень більших за найбільший елемент вибірки. Очевидно, що $\sum_{i=0}^{r+1} p_i = 1$ $\sum_{i=1}^{r} p_i = 1$. Нехай також n_i - кількість значень випадкової величини ξ, що належать множині S_i. Експериментальними частотами попадання цих значень у відповідні інтервали будуть величини n_i / n. Для великих n вони мало відрізнятимуться від ймовірностей. Але в будь-якому разі $\sum_{i=1}^{r} n_i / n = 1$. Якщо випадкова величина ξ має своїм законом розподілу якийсь відомий закон розподілу, якому відповідають вище обговорювані ймовірності p_i, то саме ці ймовірності і будуть межею частот при необмеженому зростанні обсягу вибірки. Для скінчених вибірок випадкова величина

$$\eta = \sum_{k=0}^{r+1} \frac{n}{p_k} \left(\frac{n_k}{n} - p_k \right)^2 = \sum_{k=0}^{r+1} \frac{\left(n_k - np_k \right)^2}{np_k}$$

може служити критерієм близькості теоретичної і експериментальної функцій розподілу. Структура правої частини останньої рівності відповідає структурі випадкової величини, розподіленої за законом хі-квадрат, що стає цілком зрозумілим з теореми, поданої нижче.

Нехай задана наступна випадкова величина

$$\eta = \xi_1^2 + \xi_2^2 + \ldots + \xi_r^2 ,$$

де

$$\xi_i^2 = \frac{n}{p_i}\left(\frac{n_i}{n} - p_i\right)^2 = np_i \frac{1}{p_i^2}\left(\frac{n_i}{n} - p_i\right)^2 .$$

При цьому $\frac{1}{p_i^2}\left(\frac{n_i}{n} - p_i\right)^2$ є відносною величиною квадрату відхилення експериментального значення частоти від теоретичного значення ймовірності, а np_i - вага кожного такого відхилення у загальній сумі всіх відхилень. Якщо експериментальні частот і теоретичні ймовірності описують одну і ту саму випадкову величину. то числове значення цього критерію буде меншим. Якщо різні - то більшим. І тим більшим, чим більше відрізняються експериментальні частоти від теоретичних ймовірностей. Але з чим порівнювати величину η? Як функція випадкових величин, вона так само є випадковою величиною. Можливість порівняння виникла б лише у разі, якщо функція розподілу випадкової величини η була б відома. Пірсон не лише запропонував критеріальний вираз але і довів терему про те, що при достатньо великому обсягу вибірки випадкова величина η описується розподілом ймовірності хі-квадрат з $r-1$ ступенем вільності. Якщо теоретична густина розподілу хі-квадрат для r ступенів вільності має вигляд

$$P_\xi(x) = \frac{1}{\Gamma\left(\frac{r}{2}\right)2^{\frac{r}{2}}} x^{\frac{r}{2}-1} \exp\left(-\frac{x}{2}\right),$$

то

$$P(\eta<\eta_0)=\int_0^{\eta_0}\frac{1}{\Gamma\left(\frac{r-1}{2}\right)2^{\frac{r-1}{2}}}x^{\frac{r-1}{2}-1}\exp\left(-\frac{x}{2}\right)dx\,,$$

$$P(\eta>\eta_0)=\int_{\eta_0}^{\infty}\frac{1}{\Gamma\left(\frac{r-1}{2}\right)2^{\frac{r-1}{2}-1}}x^{\frac{r-1}{2}-1}\exp\left(-\frac{x}{2}\right)dx\,.$$

Задаючи ймовірність $P(\eta<\eta_0)$, аналогічно тому, як ми задавали вище довірчу ймовірність, ми отримуємо теоретичну можливість знаходження граничного значення η_0. Використовуючи критерій Пірсона, ми отримуємо експериментальну можливість обчислення величини η. Якщо $\eta<\eta_0$, то значення величини η виявляється достатньо малим, експериментальні частоти близькі до теоретичних ймовірностей а межі нескінченно великих обсягів вибірки просто з ними збігаються і наша гіпотеза підтверджується. Якщо ж $\eta>\eta_0$ то значення величини η виявляється достатньо великими, експериментальні частоти далекі від теоретичних ймовірностей а межі нескінченно великих обсягів вибірки з ними не збігаються і наша гіпотеза не підтверджується. Точніше можна сказати так: з ймовірністю $P(\eta<\eta_0)$, яка нами задається наперед, наша гіпотеза підтверджується. З ймовірністю $P(\eta>\eta_0)$ наша гіпотеза не підтверджується. Звичайно ймовірність $P(\eta<\eta_0)$ беруть надзвичайно близькою до одиниці, так щоб спостереження значень випадкової величини, що не узгоджуються з гіпотезою, майже не зустрічались. Іншими словами, для кожної серії випробувань критеріальна випадкова величина η матиме дещо інше значення, у тому числі і таке, що не узгоджується з висунутою гіпотезою. Проте вибірки з таким значенням критерію

Пірсона зустрічатимуться вкрай рідко, якщо ймовірність $P(\eta < \eta_0)$ близька до одиниці.

Приклад 1. *В якості базового рівняння візьмемо універсальне співвідношення між довірчою ймовірністю і довірчим інтервалом для розподілу хі-квадрат*

$$P(\eta < \eta_0) := \int_0^{\eta_0} \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} x^{\frac{r}{2}-1} \exp\left(-\frac{x}{2}\right) dx .$$

Верхню межу довірчого інтервалу η_0 *підберемо такою, щоб довірча ймовірність була близькою до одиниці, наприклад* $P(\eta_0, r) = 0.997$ *. Для кожної вибірки буде своє значення* η_0*.*

За допомогою генератора випадкових чисел створимо вибірку обсягом сто елементів для нормально розподіленої випадкової величини

$$\eta := rnorm(100, 10, 2)$$

$\eta^T =$

	0	1	2	3	4	5	6	7
0	9.178	11.748	8.862	8.181	8.265	12.194	7.954	...

Упорядкуємо її у порядку зростання її елементів

$\mathrm{csort}(\eta, 0)^T =$

	0	1	2	3	4
0	4.548	5.412	6.566	6.738	...

Для створення варіаційного ряду використаємо оператор $csort(\xi, 0)$

Мінімальне і максимальне значення елементів вибірки є такими:

$$\xi_{\min} := 4.198, \qquad \xi_{\min} := 16.096 .$$

Довжина інтервалу, де знаходяться всі елементи вибірки така

$$\xi_{\max} - \xi_{\min} := 11.898 .$$

Цей інтервал зручно розбити на шість менших інтервалів однакової довжини ($r = 6$ *)*

$$h := \frac{\xi_{\max} - \xi_{\min}}{6} = 1.983 .$$

Межі цих інтервалів такі:

$$x_0 := \xi_{\min} = 4.198,\ x_1 := \xi_{\min} + h = 6.181,$$

$$x_2 := \xi_{\min} + 2h = 8.164,\ x_3 := \xi_{\min} + 3h = 10.147,$$

$$x_4 := \xi_{\min} + 4h = 12.130,\ x_5 := \xi_{\min} + 5h = 14.113,$$

$$x_6 := \xi_{\max} = 16.096 .$$

Кількості елементів вибірки, що потрапляють в кожний з малих інтервалів, такі:

$$n_1 := 4,\ n_2 := 18,\ n_3 := 36,\ n_4 := 34,\ n_5 := 6,\ n_6 := 2 .$$

Оскільки гіпотеза полягає в нормальності розподілу, то до наведених кількостей додамо також і кількості елементів вибірки, що належать інтервалам $(-\infty, x_0)$ *і* (x_6, ∞) *. Це* $n_0 := 0$, $n_7 := 0$. *Відповідні частоти:*

$$\frac{n_1}{n} := 0.04,\ \frac{n_2}{n} := 0.18,\ \frac{n_3}{n} := 0.36,\ \frac{n_4}{n} := 0.34,\ \frac{n_5}{n} := 0.06,\ \frac{n_6}{n} := 0.02 .$$

Оскільки математичне очікування нам невідоме, то в якості його наближеної оцінки візьмемо вибіркове середнє

$$\overline{\xi} := \frac{1}{n}\sum_{k=0}^{n-1} \xi_k .$$

Для виправленої вибіркової дисперсії ми використаємо вираз

$$S2(n) := \frac{1}{n-1}\sum_{k=0}^{n-1} (\xi_k - \overline{\xi})^2 .$$

Теоретичні ймовірності потрапляння елементів вибірки в менші інтервали визнаються так:

$$p_1 := \frac{1}{\sqrt{2\pi S2}} \int_{x_0}^{x_1} \exp\left(-\frac{(x - \xi\xi)^2}{2} \right) dx = 0.02 ,$$

$$p_2 := \frac{1}{\sqrt{2\pi S2}} \int_{x_1}^{x_2} \exp\left(-\frac{(x - \xi\xi)^2}{2} \right) dx = 0.129 ,$$

$$p_3 := \frac{1}{\sqrt{2\pi S2}} \int_{x_2}^{x_3} \exp\left(-\frac{(x-\xi\xi)^2}{2}\right) dx = 0.333,$$

$$p_4 := \frac{1}{\sqrt{2\pi S2}} \int_{x_3}^{x_4} \exp\left(-\frac{(x-\xi\xi)^2}{2}\right) dx = 0.346,$$

$$p_5 := \frac{1}{\sqrt{2\pi S2}} \int_{x_4}^{x_5} \exp\left(-\frac{(x-\xi\xi)^2}{2}\right) dx = 0.145,$$

$$p_6 := \frac{1}{\sqrt{2\pi S2}} \int_{x_5}^{x_6} \exp\left(-\frac{(x-\xi\xi)^2}{2}\right) dx = 0.024.$$

Якщо врахувати ймовірності для елементів вибірки набути значення менші за мінімальне і більші за максимальне

$$p_0 := \frac{1}{\sqrt{2\pi S2}} \int_{0}^{x_0} \exp\left(-\frac{(x-\overline{\xi})^2}{2}\right) dx = 0.024,$$

$$p_7 := \frac{1}{\sqrt{2\pi S2}} \int_{x_6}^{2x_6} \exp\left(-\frac{(x-\overline{\xi})^2}{2}\right) dx = 0.024,$$

то умова нормування для теоретичних ймовірностей виконуватиметься з високою точністю

$$\sum_{k=0}^{7} p_k = 1.$$

У той же час умова нормування для частот виконується точно в межах інтервалу від мінімального до максимального значень елементів вибірки

$$\sum_{k=1}^{6} \frac{n_k}{n} = 1.$$

Тепер можна обчислити числове значення критерію Пірсона

$$\sum_{k=0}^{7} \frac{(n_k - 100 p_k)^2}{100 p_k} = 4.671.$$

Очевидно, це значення у багато разів менше за верхню межу довірчого інтервалу – 19.5 і з довірчою ймовірністю 0.997 гіпотезу про нормальний розподіл вихідної випадкової величини слід прийняти. Насправді ми від початку знали, що вихідна *випадкова величина розподілена за нормальним законом. Наша критеріальна оцінка, базована на гіпотеза про нормальний розподіл, просто підтвердила цей відомий нам факт.*

Розглянемо випадок, коли гіпотеза про закон розподілу не відповідає реальному розподілу випадкової величини. Для тої самої вибірки припустимо, що закон її розподілу рівномірний. У цьому разі теоретичні ймовірності приналежності елементів вибірки до малих проміжків будуть однаковими і дорівнюватимуть 1/6. Тепер числове значення критерію Пірсона буде таким

$$\sum_{k=1}^{6}\frac{(n_k-100(1/6))^2}{100(1/6)}=184.55\,.$$

Очевидно, що у цьому разі числове значення критерію Пірсона перевищує верхню межу довірчого інтервалу і така гіпотеза з ймовірністю 0.997 має бути відкинута.

***3.10.Регресійний* аналіз**

Регресійні задачі відрізняються від задач інтерполяційних. В останніх також за результатами вимірювань будується функція, як правило поліном, така, що її значення в усіх точках $y_i, x_{i1}, x_{i2}..., x_{in}$, які є результатами вимірювань, збігаються з результатами вимірювань. Степінь інтерполяційного полінома є на одиницю меншою від кількості зазначених точок. Графік інтерполяційного полінома в точності проходить через точки $y_i, x_{i1}, x_{i2}..., x_{in}$ і дозволяє наближено знайти значення функції в усіх інших точках, що не збігаються з наведеними. Таким чином вигляд інтерполяційного полінома наперед не фіксується і залежить лише від кількості виміряних значень $y_i, x_{i1}, x_{i2}..., x_{in}$ – вузлів інтерполяції. Такий підхід доцільний у разі, якщо результати вимірювань мають високу точність і похибкою можна знехтувати.

Якщо похибкою знехтувати не можна, то вимога, щоб графік інтерполяційного полінома проходив через усі вузли інтерполяції може лише погіршити очікувану аналітичну залежність. Альтернативний підхід у випадку значної похибки, тобто регресійний підхід, полягає у тому, щоб з якихось міркувань задати вигляд очікуваної аналітичної залежності, ввівши в неї невідомі коефіцієнти. Останні можна отримати за результатами вимірювань. При цьому більша кількість вимірювань дозволяє точніше знайти введені коефіцієнти і не впливає на загальний вигляд шуканої аналітичної залежності.

Найпростішою регресією є лінійна регресія. У цьому разі ми шукаємо коефіцієнти a і b лінійної функції $y = ax + b$. Принциповим тут є те, що для кожного конкретного x величина y вимірюється із скінченою похибкою δ.

Для визначення сталих a і b проведемо n вимірювань величини y. Нехай це будуть значення $y_1, y_2 ..., y_n$, для відповідного набору значень величини x: $x_1, x_2, ..., x_n$. При цьому нехай значення $x_1, x_2, ..., x_n$ знайдені з похибкою δi, $i = 1, ..., n$.

Основним припущенням для подальшого є те, що величини δi можна вважати випадковими і незалежними. Якщо вважати їх також нормально розподіленими, то для знаходження параметрів a і b можна використати метод найбільшої правдоподібності. Оскільки

$$y_i = ax_i + b + \delta i ,$$

то y_i також є випадковою величиною з нормальним законом розподілу. Вважаючи $M\delta i = 0$, $D\delta i = \sigma^2$ для густини розподілу δi маємо

$$P\left(\delta i, a, b, \sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\delta i^2}{2\sigma^2}\right),$$

де σ^2 теж доцільно вважати невідомим параметром. Підклавши сюди $\delta i = y_i - ax_i - b$ ми отримаємо густину розподілу для випадкової величини y_i

$$P\left(y_i, a, b, \sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{\left(y_i - ax_i - b\right)^2}{2\sigma^2}\right].$$

Цей розподіл зручніший, оскільки в нього входять лише вимірювані величини x_i, y_i. Функція правдоподібності має вигляд

$$L = L\left(y_i, ..., y_n, a, b, \sigma^2\right) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left(y_i - ax_i - b\right)^2\right].$$

Тепер легко отримати і систему рівнянь правдоподібності

$$\frac{\partial \ln(L)}{\partial a} = \frac{1}{\sigma^2}\sum_{k=1}^{n} x_i\left(y_i - ax_i - b\right) = 0,$$

$$\frac{\partial \ln(L)}{\partial b} = \frac{1}{\sigma^2}\sum_{k=1}^{n}\left(y_i - ax_i - b\right) = 0,$$

$$\frac{\partial \ln(L)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{k=1}^{n}\left(y_i - ax_i - b\right)^2 = 0.$$

Розв`язання цієї системи рівнянь не становить жодних проблеми, проте найпростіший випадок виникає для $x_1 + x_2 + ... + x_n = 0$. Тоді

$$\tilde{a} = \sum_{k=1}^{n} x_k y_k \Big/ \sum_{k=1}^{n} x_k^2,$$

$$\tilde{b} = \frac{1}{n}\sum_{k=1}^{n} y_k,$$

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}\left(y_i - \tilde{a}x_i - \tilde{b}\right)^2.$$

Цим формулам можна надати і іншого вигляду, виключивши з них y_i. Тепер

$$\tilde{a} = a + \sum_{k=1}^{n} x_i \delta_i / \sum_{k=1}^{n} x_i^2 ,$$

$$\tilde{b} = b + \frac{1}{n} \sum_{k=1}^{n} \delta_i .$$

Останні дві рівності дозволяють легко перевірити незміщеність і слушність наших оцінок a і b. Дійсно, незміщеність випливає з того, що

$$M\tilde{a} = a , \; M\tilde{b} = b ,$$

Слушність з того, що

$$D\tilde{a} = \sigma^2 / \sum_{k=1}^{n} x_k^2 , \; D\tilde{b} = \sigma^2 / n ,$$

а тому

$$\lim_{n \to \infty} D\tilde{a} = \lim_{n \to \infty} D\tilde{b} = 0 .$$

Якщо припущення про нормальність розподілу δi не відповідає дійсності, то оцінки параметрів a і b можна отримати з умови мінімуму виразу

$$Q(a,b) = \sum_{k=1}^{n} \left(y_k - ax_k - b \right)^2 = \sum_{k=1}^{n} \delta_k^2 ,$$

тобто з умови мінімального значення суми квадратів відхилень шуканої прямої від експериментальних точок. Цей метод називається методом найменших квадратів. Тут ми його застосуємо для відтворення лінійної залежності. Знайдемо цим методом оцінки параметрів a і b. Очевидно оцінки будуть екстремумами функції $Q(a,b)$

$$\frac{\partial Q(a,b)}{\partial a} = 2\sum_{k=1}^{n} x_k \left(y_k - ax_k - b \right) = 0 ,$$

$$\frac{\partial Q(a,b)}{\partial b} = 2\sum_{k=1}^{n}\left(y_k - ax_k - b\right) = 0 .$$

Отримані рівняння збігаються з відповідними рівняннями методу найбільшої правдоподібності і у разі, якщо $x_1 + x_2 + ... + x_n = 0$ дають вже наведений вище результат

$$\tilde{a} = \sum_{k=1}^{n} x_k y_k \Big/ \sum_{k=1}^{n} x_k^2 ,$$

$$\tilde{b} = \frac{1}{n}\sum_{k=1}^{n} y_k .$$

Вони, очевидно, є незміщеними і слушними оцінками цих параметрів.

Отже, якщо похибки δi розподілені за нормальним законом, то оцінки, отримані методами найбільшої правдоподібності і найменших квадратів збігається. Якщо ж δi мають інший закон розподілу, то метод найбільшої правдоподібності дасть інший результат. Результати ж методу найменших квадратів не залежать від закону розподілу випадкових величин δi і завжди ті самі. З сказаного можна зробити висновок, що для відомої функції розподілу краще застосовувати метод найбільшої правдоподібності. Для невідомої функції розподілу – метод найменших квадратів, оскільки останній дає хоч якусь оцінку параметрів.

Для довільної суми величин $x_1, x_1, ..., x_n$

$$\tilde{a} = \frac{n\sum_{k=1}^{n} x_k y_k - \sum_{k=1}^{n} x_k \sum_{k=1}^{n} y_k}{n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2},$$

$$\tilde{b}=\frac{\sum\limits_{k=1}^{n}x_k^2\sum\limits_{k=1}^{n}y_k-\sum\limits_{k=1}^{n}x_k\sum\limits_{k=1}^{n}x_ky_k}{n\sum\limits_{k=1}^{n}x_k^2-\left(\sum\limits_{k=1}^{n}x_k\right)^2}.$$

Приклад 1. *Побудуємо лінійну регресію для наступних даних, що є натуральними числами на відрізку* $[1,2,\ldots,21]$

$x^T =$

	0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9	...

.

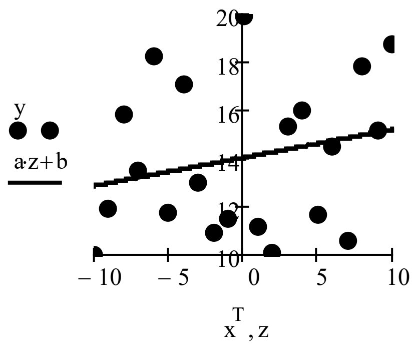
$$\Delta:=\begin{vmatrix}\sum\limits_{k=0}^{20}x_k^2 & \sum\limits_{k=0}^{20}x_k\\ \sum\limits_{k=0}^{20}x_k & 21\end{vmatrix},\ \Delta_a:=\begin{vmatrix}\sum\limits_{k=0}^{20}x_ky_k & \sum\limits_{k=0}^{20}x_k\\ \sum\limits_{k=0}^{20}y_k & 21\end{vmatrix},\ \Delta_b:=\begin{vmatrix}\sum\limits_{k=0}^{20}x_k^2 & \sum\limits_{k=0}^{20}x_ky_k\\ \sum\limits_{k=0}^{20}x_k & \sum\limits_{k=0}^{20}y_k\end{vmatrix}.$$

Відповідно,

$$a:=\frac{\Delta_a}{\Delta},\ b:=\frac{\Delta_b}{\Delta}.$$

На графіку це виглядатиме так

$$z:=1,1.01..21$$



Мал. 1. *Експериментальні точки і лінійна регресія.*

Тут випадковий вектор

$$y := runif(21,10,20)$$

Має рівномірно розподілені 21 компоненту на відрізку $[10,20]$. *З малюнка видно, що лінійна регресія в хаотичному нагромадженні точок формує тенденцію до лінійного зростання.*

3.11.Метод найменших квадратів

Нехай ми експериментально досліджуємо функційну залежність $y = f(x)$, не виходячи з припущення про її лінійність, як у попередньому параграфі. Для цього ми задаємо певну послідовність $\{x_i\}$ значень величини x і вимірюємо відповідну послідовність $\{y_i\}$ значень величини y. Задача полягає в аналітичному представленні шуканої функційної залежності. Надалі розглянемо той випадок, коли результати вимірювання містять значну помилку, зумовлену будь – якими причинами. Тоді втрачає сенс намагання точно описати експериментальні дані. Іншими словами, графік функції $y = f(x)$ не повинен проходити через всі точки з координатами $\{x_i, y_i\}$. Швидше він повинен передавати певну тенденцію у розташуванні точок. Більше того, спроба побудувати інтерполяційний поліном, тобто поліном, графік якого проходить через всі точки, лише погіршила δ ситуацію. У цьому разі ми могли б отримати абсолютно нефізичний результат. Питання про розмежування підходів, основаних на побудові інтерполяційних поліномів і емпіричних формул, не є простим.

Емпіричні функції можуть належати до довільного класу функцій. Позначимо цю функційну залежність так

$$y = f\left(x; a_0, a_1, ..., a_n\right),$$

де $a_0, a_1, ..., a_n$ – параметри, які підлягають визначенню. При побудові інтерполяційного полінома кількість точок $\{x_i, y_i\}$ збігається з кількістю параметрів $\{a_i\}$. У методі найменших квадратів кількість точок може значно перевищувати кількість коефіцієнтів. Тим самим, недостатня точність координат окремих точок компенсується їх більшою кількістю. Можна довести, що оцінка параметрів $\{a_i\}$ методом найменших квадратів є незміщеною і слушною.

Суть метода найменших квадратів полягає у наступній умові для знаходження параметрів $\{a_i\}$. Нехай

$$S = \sum_{k=1}^{N} \left[y_k - f\left(x_k; a_0, a_1, ..., a_n\right)\right]^2$$

- сума квадратів відхилень виміряної фізичної величини y від розрахованих $f\left(x; a_0, a_1, ..., a_n\right)$ приймає мінімальне значення. Наведена формула вірна для набору експериментальних даних, виміряних з однаковою точністю (дисперсією). Якщо всі вимірювані результати мають різну точність (дисперсію σ_i), то остання сума заміниться наступною

$$S = \sum_{k=1}^{N} \left[y_k - f\left(x_k; a_0, a_1, ..., a_n\right)\right] / \sigma_k^2 .$$

З умови екстремуму цієї функції ми і отримаємо систему рівнянь для знаходження невідомих параметрів

$$\frac{\partial S(a_0, a_1, ..., a_n)}{\partial a_0} = 0,$$

$$\frac{\partial S(a_0, a_1, ..., a_n)}{\partial a_1} = 0,$$

……………………

$$\frac{\partial S(a_0, a_1, ..., a_n)}{\partial a_n} = 0 .$$

Ясно, що в загальному випадку цю систему рівнянь можна розв'язати лише у числовий спосіб.

3.12.Задачі для самостійного розв`язання

Варіант 1.

1.Маємо 4 заготівки для виготовлення деталей. Ймовірність виготовлення придатної деталі дорівнює 0.75. Знайти закон розподілу випадкової величини, що є кількістю заготівок, які буде використано для виготовлення придатної деталі. Знайти математичне очікування і дисперсію випадкової величини, а також імовірність того, що із цих заготовок буде виготовлена хоча б одна стандартна деталь.

2.Знайти математичне очікування випадкової величини, що має розподіл Пуассона.

3.У цеху встановлені 5 верстатів. Протягом 25 днів реєструвалась кількість верстатів, які не працювали. Здобуті такі значення: 0, 1, 2, 1, 1, 2, 3, 2, 1, 4, 2, 0, 0, 2, 2, 3, 3, 1, 0, 1, 2, 1, 3, 5, 0. Побудувати статистичну функцію розподілу. Обчислити вибіркові середнє та дисперсію.

4.Вибірку обсягом n зроблено із генеральної сукупності, розподіленої за законом Релея

$$P_\xi(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), x > 0 .$$

Знайти точкову оцінку для параметра σ^2.

5.За даними вибірки обсягом n з нормально розподіленої генеральної сукупності, дисперсія якої відома, а довірча ймовірність дорівнює γ, знайти довірчий інтервал для математичного сподівання.

Варіант 2.

1.Задано функцію

$$F_\xi(x)=\begin{cases}0, x<2.5,\\ ax^2+bx, 2.5\le x\le 4,\\ 1, x>4\end{cases}$$

Довести, що можна підібрати такі значення a і b, для яких функція $F_\xi(x)$ буде функцією розподілу ймовірностей випадкової величини ξ. Знайти $P(3\le\xi\le 4)$.

2.Знайти середнє значення кінетичної енергії молекул ідеального газу, що мають розподіл Максвела.

3.Із генеральної сукупності зроблено вибірку обсягом 32. Здобуті такі реалізації випадкової величини: 2.2, 7.1, 6.3, 3.9, 5.9, 5.6, 4.7, 7.9, 3.2, 6.1, 5.5, 6.4, 6.0, 6.9, 4.7, 6.4, 6.9, 6.7, 7.9, 4.2, 6.7, 6.0, 9.2, 5.5, 6.5, 3.5, 4.9, 7.2, 4.9, 8.9, 5.7. Скласти варіаційний ряд і побудувати гістограму. Запропонувати гіпотезу про вигляд функції розподілу ймовірності для даної випадкової величини. Знайти вибіркові середнє, дисперсію, асиметрію і ексцес.

4.Методом моментів за даними вибірки обсягом n знайти оцінку параметра p геометричного розподілу

$$P(m)=p(1-p)^{m-1},\quad m\in N.$$

5.За даними вибірки обсягом n з нормально розподіленої генеральної сукупності, дисперсія якої невідома, а довірча ймовірність дорівнює γ, знайти довірчий інтервал для математичного сподівання.

Варіант 3.

1.Задано функцію $F_\xi(x)=a\sin(x)+b\cos(x)+0.5$. Довести, що коли $x\in[0,\pi/2]$, то можна знайти a і b, такі що $F_\xi(x)$ є функцією розподілу ймовірностей. Знайти ці значення та густину ро-поділу $P_\xi(x)$.

2.Знайти середнє значення абсолютної величини швидкості молекул ідеального газу, що мають розподіл Максвела.

3.Кількість деталей, потрібних для ремонту обладнання на тиждень, визначалася на підставі спостережень, здійснюваних протягом 20 тижнів. У результаті були здобуті такі значення: 0, 1, 1, 1, 0, 0, 2, 3, 0, 0, 1, 1, 4, 0, 5, 2, 3, 4, 2, 3. Побудувати статистичну функцію розподілу. Обчислити вибіркові середнє і дисперсію і порівняти їх з математичним очікуванням і дисперсією, згідно з гіпотезою про закон розподілу Пуасона з параметром 1 у генеральній сукупності.

4.Із нормально розподіленої сукупності з $D\xi = \sigma^2$ зроблено вибірку обсягом n. Знайти оцінку для $M\xi$.

5.За даними вибірки обсягом n з нормально розподіленої генеральної сукупності математичне очікування якої відоме, а довірча ймовірність дорівнює γ, знайти довірчий інтервал для дисперсії.

Варіант 4.

1.Випадкова величина задається функцією розподілу

$$F_\xi(x) = \begin{cases} 0, if\ x \le 1; \\ a + b\arctan(x), if\ x > 1. \end{cases}$$

Знайти a, b і густину розподілу $P_\xi(x)$.

2.Знайти математичне очікування випадкової величини, що має розподіл Бернулі.

3.Для оцінювання ймовірності настання події було проведено 10 серій послідовних випробувань до першого успішного випробування. В результаті здобуті такі значення: 4, 2, 5, 3, 4, 4, 3, 5, 3, 4. Побудувати статистичну функцію розподілу ймовірностей. Припускаючи, що реалізується геометричний розподіл з $p = 0.25$, знайти вибіркові середнє і дисперсію, а також математичне очікування і дисперсію для геометричного розподілу. Результати порівняти.

4.Із нормально розподіленої сукупності з $M\xi = a$ зроблено вибірку обсягом n. Знайти оцінку для $D\xi$.

5.За даними вибірки обсягом n з нормально розподіленої генеральної сукупності, математичне очікування якої відоме, а довірча ймовірність дорівнює γ, знайти довірчий інтервал для дисперсії.

Варіант 5.

1.Випадкова величина задається функцією розподілу

$$F_{\xi}(x) = \begin{cases} 0, x < a, \\ ab^2 x - 1/3, a \le x \le 2, \\ 1, x > 2 \end{cases}.$$

Знайти a, b і густину розподілу $P_{\xi}(x)$.

2.Знайти дисперсію випадкової величини, що має розподіл Коші.

3.У вимірювальному приладі встановлені декілька однотипних опорів. Під час експлуатації 15 приладів протягом року кількість опорів, які довелося замінити, була такою: 1, 3, 2, 0, 4, 1, 5, 5, 5, 4, 3, 4, 2, 1, 2. Побудувати статистичну функцію розподілу. Знайти вибіркові середнє та дисперсію і порівняти їхні значення з числовими характеристиками відповідного розподілу Пуассона з параметром 3.

4.У сукупності, розподіл у якій задається густиною ймовірності

$$P_{\xi}(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(x) - \mu)^2}{2\sigma^2} \right), x > 0,$$

зроблено вибірку обсягом n. Методом найбільшої правдоподібності знайти точкову оцінку для μ (значення σ^2 відоме).

5.За даними вибірки обсягом n з нормально розподіленої генеральної сукупності, дисперсія якої невідома, а довірча ймовірність дорівнює γ, знайти довірчий інтервал для математичного сподівання.

Варіант 6.

1.Графік густини розподілу - півколо з центром у початку координат. Знайти аналітичний вираз для густини розподілу та математичного очікування.

2.Знайти математичне очікування випадкової величини, що має розподіл Коші.

3.Маємо дані про термін служби радіоламп (у тисячах годин): 0.45, 0.21, 0.14, 0.15, 1.52, 0.1, 0.52, 1.59, 3.38, 2.25, 0.8, 1.26, 2.31, 0.84, 3.72, 2.11, 1.02, 4.2, 2.53, 0.78, 2.92, 0.71, 4.7, 3.02, 1.58, 4.12, 2.59, 0.88, 0.96, 1.76, 1.96, 1.76, 1.93, 4.9, 2.82, 1.14, 5.7, 1.21, 1.47, 3.52, 0.36, 0.64. Скласти варіаційний ряд і побудувати гістограму. Висунути гіпотезу про закон розподілу в сукупності. Знайти вибіркові середнє і дисперсію.

4.Із сукупності із густиною ймовірності

$$P_{\xi}(x)=\begin{cases} a\exp\left(-a(x-\mu)\right)^{2}, if\ x\geq\mu, \\ 0, if\ x<\mu \end{cases}$$

зроблено вибірку обсягом n. Методом найбільшої правдоподібності знайти оцінку для a і μ.

5.Як оцінку відстані до навігаційного знаку беруть середнє арифметичне незалежних вимірювань, що їх виконують n дальномірів. Похибки вимірювання окремими дальномірами розподілені нормально з нульовим математичним очікуванням і середньо квадратичним відхиленням, рівним 10 м. Скільки потрібно дальномірів, щоб абсолютна величина похибка всієї серії вимірювань ($\Delta\sigma/\sqrt{n}$) з довірчою ймовірністю 0.96 не перевищувала 15 м.

Варіант 7.

1.Випадкова величина набуває ненульових значень на відрізку $[a,2]$, де задано густину розподілу $P_{\xi}(x)=Ax^{4}$. Визначити $a, A, F_{\xi}(x), D(x)$, якщо $M\xi=0$.

2.Знайти дисперсію випадкової величини, що має рівномірний розподіл.

3.Із генеральної сукупності зроблено вибірку обсягом 30. Отримані такі вибіркові значення: 4, 4.3, 5.68, 6.2, 5.64, 5.8, 4.25, 5.4, 5.3, 5.2, 4.55, 5.32, 6, 6.15, 4.56, 6.64, 6.5, 4.7, 6.8, 6.15, 5.6, 5.1, 4.2, 4.8, 6.9, 7, 4.8, 5, 5.25, 6.2. Скласти варіаційний ряд і побудувати гістограму. Висунути гіпотезу про закон розподілу у вибірці. Обчислити вибіркові середнє і дисперсію.

4.Із сукупності із густиною ймовірності

$$P_{\xi}(x)=\begin{cases} \frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln(x)-\mu)^{2}}{2\sigma^{2}}\right), if\ x>0, \\ 0, if\ x\leq 0 \end{cases}$$

зроблено вибірку обсягом n. Методом найбільшої правдоподібності знайти точкову оцінку для σ^2 і μ.

5.У висліді вимірювання ємності 20 конденсаторів дістали для вибіркового середнього значення 4.47 пф, а для вибіркової дисперсії 0.0121 пф^2. Порівняйте точність оцінки математичного сподівання при довірчій ймовірності 0.95 використовуючи нормальний закон розподілу і розподіл Стьюдента.

Варіант 8.

1.Випадкова величина набуває ненульових значень на відрізку $[a,2]$, де задано густину розподілу $P_\xi(x) = Ax^2$. Визначити $a, A, F_\xi(x), D(x)$, якщо $M\xi = 0$.

2.Знайти математичне очікування випадкової величини, що має рівномірний розподіл.

3.У цеху встановлені 6 верстатів. Протягом 25 днів реєструвалась кількість верстатів, які не працювали. Здобуті такі значення: 0, 1, 2, 1, 6, 2, 3, 2, 1, 4, 2, 0, 6, 2, 2, 3, 3, 1, 0, 1, 2, 1, 3, 5, 0. Побудувати статистичну функцію розподілу. Обчислити вибіркові середнє та дисперсію.

4.Із генеральної сукупності із густиною ймовірності

$$P_\xi(x) = \frac{\alpha^p x^{p-1} \exp(-\alpha x)}{\Gamma(p)}, x > 0$$

зроблено вибірку обсягом n. Знайти точкову оцінку для параметра α.

5.Під час перевірки 400 лампочок середній термін їх горіння становив 1220 год. Оцінити з довірчою ймовірністю 0.95 математичне сподівання тривалості горіння, якщо середнє квадратичне відхилення дорівнює 35 год., а закон розподілу вибіркового середнього можна вважати нормальним.

Варіант 9.

1.Випадкова величина набуває ненульових значень на відрізку $[a,b]$, де задано густину розподілу $P_\xi(x) = \frac{3}{16}(x-c)^2$. З умови нормування визначити c.

2.Знайти дисперсію випадкової величини, що має нормальний розподіл.

3.Із генеральної сукупності зроблено вибірку обсягом 30. Здобуті такі реалізації випадкової величини: 2.3, 7.0, 6.5, 4.9, 3.9, 6.6, 4.7, 7.9, 3.2, 6.1, 5.5, 6.4, 6.0, 6.9, 4.7, 6.4, 6.9, 6.7, 7.9, 4.2, 6.7, 6.0, 9.2, 5.5, 6.5, 3.5, 4.9, 7.2, 4.9. Скласти варіаційний ряд і побудувати гістограму. Запропонувати гіпотезу про вигляд функції розподілу ймовірності для сукупності. Знайти вибіркові середнє, дисперсію, асиметрію і ексцес.

4.Із сукупності, розподіленої за показниковим законом

$$P_\xi(x) = \lambda \exp\left(-\lambda x\right), \quad x \ge 0$$

зроблено вибірку обсягом n. Методами моментів і найбільшої правдоподібності знайти точкову оцінку для параметра λ.

5.На основі 100 спостережень визначено, що в середньому для виробництва деталі потрібно 5.5 сек., а вибіркова дисперсія дорівнює 2.89. Вважаючи, що тривалість виготовлення деталі розподілена нормально, знайти інтервальні оцінки для математичного очікування і дисперсії з довірчою ймовірністю 0.96 і 0.98 відповідно.

Варіант 10.

1.Випадкова величина набуває значень на відрізку $[1,b]$, де задано густину розподілу $P_\xi(x) = A\ln(x)$. Визначити A, якщо $M\xi = \dfrac{e^2+1}{4}$.

2.Знайти математичне очікування випадкової величини, що має нормальний розподіл.

3.Кількість деталей, потрібних для ремонту обладнання на тиждень, визначалася на підставі спостережень, здійснюваних протягом 15 тижнів. У результаті були здобуті такі значення: 0, 1, 1, 1, 0, 0, 2, 3, 0, 0, 1, 1, 4, 0, 5. Побудувати емпіричну функцію розподілу. Вважаючи, що кількість використовуваних деталей має розподіл Пуассона з параметром розподілу, рівним 1, обчислити вибіркові середнє і дисперсію і порівняти їх з математичним очікуванням і дисперсією, згідно з висунутою гіпотезою про закон розподілу у сукупності.

4.Із нормально розподіленої сукупності

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

з $M\xi = a$ зроблено вибірку обсягом n. Знайти оцінку для $D\xi$. Перевірити її на незміщеність ефективність та обґрунтованість.

5.За даними вибірки обсягом n з нормально розподіленої генеральної сукупності математичне очікування якої відоме, а довірча ймовірність дорівнює γ, знайти довірчий інтервал для дисперсії.

Глава 4. Випадкові процеси

4.1.Випадкові процеси

Основним об'єктом математичної статистики є випадкова величина. Подальшим ускладненням об'єкту під назвою випадкова величина є випадковий процес. Якщо випадкова величина є числовою функцією, областю визначення якої є простір елементарних подій, то випадковий процес це набір таких функцій. Це набір може бути скінченим або нескінченим. Причому нескінченість можлива як злічена, так і континуальна. Якщо випадкову величину можна записати у вигляді $\xi(\omega)$, де ω - елементарна подія з простору елементарних подій, то набір випадкових величин, або випадковий процес можна задати у вигляді $\xi_t(\omega)$. Тут параметр t перераховує всі елементи множини випадкових величин. Цей параметр звичайно має сенс часу і, як вже зазначалось вище, може набувати скінченого, зліченого або нескінченого континуального набору значень. Якщо параметр t фіксований, то ми отримуємо звичайну випадкову величину, яка у даному разі називається перерізом випадкового процесу. Якщо фіксованою є елементарна подія, а параметр t змінюється, то відповідна функція називається траєкторією випадкового процесу.

Для опису перерізу випадкового процесу достатньо ($t = t_0$) функції розподілу ймовірностей однієї випадкової величини

$$F_\xi(x_0, t_0) = P(\xi_{t_0} < x_0).$$

Якщо параметр t набуває дискретного скінченого набору значень, $0 \le t_1 < t_2 < ... < t_n < \infty$, $n \in N$, випадковий процес еквівалентний заданню n випадкових величин $\xi_{t_1}, \xi_{t_2}, ..., \xi_{t_n}$. Очевидно, що з максимальною повнотою такий випадковий процес можна описати функцією розподілу ймовірностей n випадкових величин

$$F_{\xi_1, \xi_2, ..., \xi_n}(x_1, x_2, ..., x_n) = P(\xi_{t_1} < x_1, \xi_{t_2} < x_2, ..., \xi_{t_n} < x_n).$$

Якщо параметр t набуває дискретного зліченого набору значень, $0 \le t_1 < t_2 < ... < t_n ... < \infty$, випадковий процесс еквівалентний заданню нескінченого, але зліченого набору випадкових величин $\xi_{t_1}, \xi_{t_2}, ..., \xi_{t_n}, ...$. Очевидно, що з максимальною повнотою такий випадковий процес можна описати функцією розподілу ймовірностей нескінченої, але зліченої кількості випадкових величин

$$F_{\xi_1, \xi_2, ..., \xi_n}(x_1, x_2, ..., x_n, ...) = P(\xi_{t_1} < x_1, \xi_{t_2} < x_2, ..., \xi_{t_n} < x_n, ...).$$

Якщо параметр t набуває неперервного набору значень, наприклад, $t \in [0, T]$, то навіть формально, строго кажучи, такий випадковий процес з максимально можливою повнотою не можна описати функцією навіть зліченої кількості випадкових величин. Адекватним математичним апаратом тут є функційне інтегрування. Проте, функція розподілу зліченої кількості випадкових величин дозволяє досягти хоча і наближеного, але з наперед заданою точністю опису випадкового процесу. Якраз в останньому випадку специфіка дослідження випадкових процесів виявляється у найбільшій мірі порівняно з дослідженням окремих випадкових величин.

Випадкові процеси можуть набувати дискретного і неперервного набору значень. Обидва варіанти випадкових процесів знаходять широке використання на практиці. До найпопулярнішого випадкового процесу з дискретним набором значень належить процес Пуасона. Тут значення випадкової функції (випадкового процесу) змінюються у часі стрибком. Відповідно, до найпопулярнішого випадкового процесу з неперервним набором значень належить процес Вінера. Ці випадкові процеси ми і розглянемо нижче.

4.2.Пуасонівський процес

Пуасонівський процес настільки важливий для практичного використання, що ми розглянемо два способи його отримання: оснований на фізичних міркуваннях і на використанні базових засад теорії ймовірностей і математичної статистики. Випадковий процес Пуасона

Нехай нас цікавить ймовірність того, що за проміжок часу $[0,t]$ відбудеться r подій, що задовольняють наступним умовам:

1. Всі події є незалежними одна від одної.
2. Ймовірність однієї події $P_1(dt)$ за малий проміжок часу dt прямо пропорційна цьому проміжку часу $P_1(dt) = \mu dt$.
3. Ймовірність відсутності події за проміжок часу dt $P_0(dt) = 1 - \mu dt$.
4. Ймовірність декількох подій за проміжок часу dt дорівнює нулю $P_2(dt) = ... = P_r(dt) = 0$.

Наведених умов достатньо для знаходження закону розподілу дискретної випадкової величини r. В якості простору елементарних подій візьмемо множину точок проміжку $[0,\infty)$. В цьому просторі елементарних подій розглянемо дві незалежні події $[0,t]$ і $[t,t+dt]$. Розглянемо випадкову функцію – ймовірність того, що випадкова величина r, визначена на даному просторі елементарних подій, дорівнює нулю, тобто функцію $P_0(t+dt)$. Аргумент цієї функції можна виразити через введені вище події, які мають відбутись одночасно $P_0(t+dt) = P_0([0,t][t,t+dt])$. Оскільки події незалежні, то за теоремою множення ймовірностей маємо

$$P_0(t+dt) = P_0(t)P_0(dt) = P_0(t)(1-\mu dt).$$

Останнє рівняння можна записати у вигляді

$$P_0(t+dt) - P_0(t) = dP_0(t) = -\mu P_0(t)dt$$

або у вигляді диференційного рівняння

$$\frac{dP_0(t)}{dt} + \mu P_0(t) = 0.$$

Загальним розв'язком цього рівняння є

$$P_0(t) = A\exp(-\mu t)\,.$$

Довільну сталу можна знайти з очевидної початкової умови $P_0(0) = 1$, яка означає, що у початковий момент часу не могло відбутись жодної події з ймовірністю 1. Отже, шуканий частинний розв’язок матиме вигляд

$$P_0(t) = \exp(-\mu t)\,.$$

Отже ймовірність того, що за проміжок часу $[0,t]$ не відбудеться жодної події, експоненційно спадає до нуля із зростанням цього проміжку часу.

Нехай тепер за проміжок часу $[0,t+dt]$ відбудеться лише одна подія $r=1$. Тут можливі дві гіпотези: ця подія відбулась за проміжок часу $[0,t]$ і не відбулась за проміжок часу $[t,t+dt]$ та ця подія відбулась за проміжок часу $[t,t+dt]$ і не відбулась за проміжок $[0,t]$. Ймовірність реалізації першої гіпотези є $P_1(t)P_0(dt)$. Ймовірність реалізації другої гіпотези - $P_0(t)P_1(dt)$. Відповідно до формули повної ймовірності

$$P_1(t+dt) = P_1(t)P_0(dt) + P_0(t)P_1(dt)\,,$$

або

$$P_1(t+dt) = P_1(t)(1-\mu dt) + P_0(t)\mu dt\,.$$

Це рівняння зручно записати так

$$\frac{dP_1(t)}{dt} + \mu P_1(t) = \mu P_0(t)$$

або

$$\frac{dP_1(t)}{dt} + \mu P_1(t) = \mu \exp(-\mu t)\,.$$

Тут початкова умова також очевидна, це $P_1(0) = 0$, у початковий момент ймовірність відбутись одній події не може бути відмінною від нуля, оскільки це протирічить раніше виписаній початковій умові $P_0(0)=1$. Можна показати, що шуканим частинним розв'язком цього рівняння є

$$P_1(t)=\mu t\exp(-\mu t).$$

Дійсно, загальний розв'язок однорідного рівняння має вигляд

$$P_1(t)=A\exp(-\mu t).$$

Шукатимемо частинний розв'язок неоднорідного рівняння методом варіації довільної сталої, тобто у вигляді

$$P_1(t)=A(t)\exp(-\mu t).$$

Після підстановки його у вихідне рівняння, отримаємо

$$\frac{dA(t)}{dt}=\mu.$$

Звідси

$$A(t)=\mu t,$$

а частинний розв'язок неоднорідного рівняння матиме вигляд

$$P_1(t)=\mu t\exp(-\mu t).$$

Загальний розв'язок неоднорідного рівняння

$$P_1(t)=A\exp(-\mu t)+\mu t\exp(-\mu t).$$

Початкова умова для шуканої функції дасть значення довільної сталої $A=0$ *і матимемо той самий частинний розв'язок неоднорідного рівняння*

$$P_1(t)=\mu t\exp(-\mu t).$$

Рівняння для ймовірності того, що за проміжок часу $[0,t+dt]$ *відбудеться дві події для* $P_2(t)$ *матиме вигляд*

$$\frac{dP_2(t)}{dt}+\mu P_2(t)=\mu^2 t\exp(-\mu t).$$

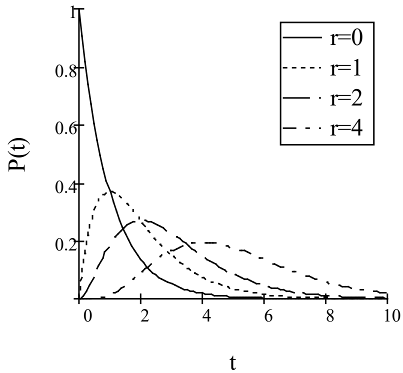
Розв'язавши його аналогічно попередньому, матимемо

$$P_2(t)=\frac{\mu^2 t^2}{2!}\exp(-\mu t).$$

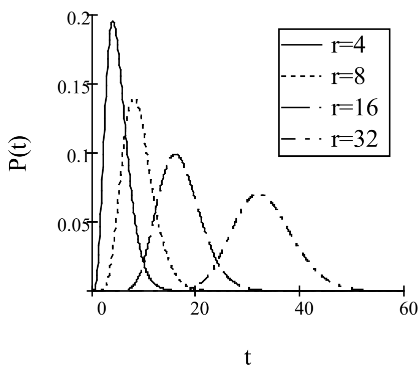
Продовжуючи цей процес отримання і розв'язання рівнянь до розгляду за проміжок часу $[0, t+dt]$ r *подій, отримаємо*

$$P_r(t) = \frac{\mu^r t^r}{r!} \exp(-\mu t).$$

Наступний малюнок показує вигляд розподілу Пуассона для різних значень випадкової величини



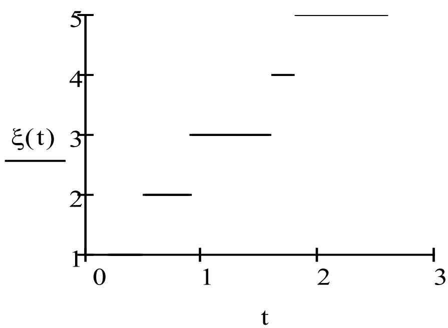
Малюнок 1. *Вигляд розподілу Пуасона для малих значень випадкової величини.*



Малюнок 2. *Вигляд розподілу Пуасона для великих значень випадкової величини.*

З малюнків видно, що для малих часів ймовірність того, що не відбудеться жодної події, близька до нуля і поступово з часом зменшується. Одночасно зростає ймовірність того, за час t відбудеться одна подія. Далі зменшується і ця ймовірність, зате зростає ймовірність того, що відбудуться дві події тощо. Крім того, при досягненні кількістю подій певної, не дуже великої величини ймовірність для випадкового процесу Пуасона швидко наближається до нормального розподілу. Коли $r = 32$ ці розподіли візуально не розрізняються.

Траєкторія випадкового процесу Пуасона має вигляд



Малюнок 3. *Траєкторія випадкового процесу Пуасона.*

Перейдемо до фундаментальнішого дослідження випадкового процесу Пуасона. Розглянемо декілька базових для теорії випадкових процесів означень.

Випадковий процес з неперервний часом ξ_t, для простоти ми не пишемо його аргумент, що належить простору елементарних подій, називається процесом з незалежними приростами, якщо для довільних $0 \le t_1 < t_2 < ... < t_n < \infty$, випадкові величини ξ_{t_1}, $\xi_{t_2} - \xi_{t_1}$, … , $\xi_{t_n} - \xi_{t_{n-1}}$ є незалежними. Це означає, що різниця між сусідніми членами послідовності випадкових величин ніяк не залежить від різниці будь-якої іншої пари сусідніх членів цієї ю послідовності.

Випадковий процес називається однорідним, якщо різниця випадкових величин $\xi_{t_k} - \xi_{t_{k-1}}$ і $\xi_{t_k+h} - \xi_{t_{k-1}+h}$, $k = 1,...,n$ однаково розподілені. Це означає, що зсув у часі кожної з випадкових величин не змінює функцію розподілу ймовірностей їх різниці.

Пуасонівським процесом називається процес, що задовольняє наступним умовам:

1. ξ_t - випадковий процес з незалежними приростами.
2. ξ_t - випадковий процес, однорідний у часі.
3. $\xi_{t=0}(\omega) = 0$.

Фактично, остання умова є початковою умовою процесу. Це означає, що у початковий момент часу переріз випадкового процесу дорівнює нулю, тобто всім елементам простору елементарних подій ставиться у відповідність цілком конкретне і однакове для всіх нульове значення. Мовою ймовірності це означає, що у початковий момент часу з ймовірністю $P(\xi_{t=0} = 0) = 1$ нічого не відбувається.

4. $P(\xi_{t=h} = 0) \approx 1 - \mu h$, $P(\xi_{t=h} = 1) \approx \mu h$, $P(\xi_{t=h} = 2) \approx 0$.

Останні умови визначають характер випадкового процесу невдовзі після його початку, коли $h << 1$. Перша умова означає, що при віддаленні від початкового моменту часу ймовірність того, що нічого не відбудеться, зменшується за лінійним законом. Тут μ вважається відомою сталою. Друга умова означає, що при віддаленні від початкового моменту часу ймовірність того, що відбудеться одна подія (значення випадкової величини дорівнюватиме 1) лінійно зростає. Третя умова означає, що при віддаленні від початкового процесу ймовірністю того, що відбудуться дві події, можна знехтувати.

Третя і четверта умови задають значення випадкового процесу у початковий момент часу і цілком визначають його поведінку в безпосередньому околі початкового моменту. Перша і друга умови носять загальний характер і разом з третьою і четвертою дозволяють визначити випадковий процес Пуасона у довільний момент часу.

Функція розподілу процесу Пуасона має вигляд

$$P(\xi_t = k) = \frac{\mu^k t^k}{k!}\exp(-\mu t)\,,\ k = 0,1,2,\ldots.$$

Доведення. *Доведення проведемо на основі аналізу твірної функції розподілу випадкової величини* ξ_t. *Якщо випадкова величина набуває лише цілих невід'ємних значень, то твірна функція її розподілу має вигляд наступного математичного очікування* $\varphi_t(x) = M(x^{\xi_t})$. *Надамо параметру* t *малий приріст* h. *Тоді*

$$\varphi_{t+h}(x) = M(x^{\xi_{t+h}}) = M(x^{\xi_{t+h}-\xi_t+\xi_t}) = M(x^{\xi_{t+h}-\xi_t}x^{\xi_t}).$$

Оскільки випадковий процес ξ_t *є випадковим процес з незалежними приростами, тобто* $\xi_{t+h}-\xi_t$ *і* ξ_t *є незалежними величинами, то за властивістю математичного очікування добутку незалежних величин*

$$\varphi_{t+h}(x) = M(x^{\xi_{t+h}-\xi_t}x^{\xi_t}) = M(x^{\xi_{t+h}-\xi_t})M(x^{\xi_t}) = \varphi_t(x)M(x^{\xi_{t+h}-\xi_t}).$$

Далі скористаємось властивістю однорідності у часі даного випадкового процесу. Це означає, що випадкові величини $\xi_{t+h}-\xi_t$ *і* $\xi_h-\xi_{t=0}$ *однаково розподілені, а отже мають і однакові твірні функції* $M(x^{\xi_{t+h}-\xi_t}) = M(x^{\xi_h-\xi_{t=0}})$. *Далі врахуємо початкову умову для випадкової величини* $\xi_{t=0}=0$. *Тоді* $M(x^{\xi_{t+h}-\xi_t}) = M(x^{\xi_h})$ *Нарешті, для малих* $h \ll 1$ *останнє математичне очікування можна обчислити безпосередньо*

$$M(x^{\xi_h}) \approx x^{\xi_h=0}P(\xi_{t=h}=0)+x^{\xi_h=1}P(\xi_{t=h}=1)=1-\mu h+x\mu h,\ |x|\le 1.$$

Таким чином,

$$\varphi_{t+h}(x)=\varphi_t(x)M(x^{\xi_{t+h}-\xi_t})=\varphi_t(x)M(x^{\xi_h})\approx\varphi_t(x)(1-\mu h+x\mu h),$$

або

$$\frac{\varphi_{t+h}(x)-\varphi_t(x)}{h}\approx\mu(x-1)\varphi_t(x).$$

Якщо перейти до межі, то остання наближена рівність перетворить у строгу рівність, а різницеве рівняння перетвориться у диференційне

$$\frac{d\varphi_t(x)}{dt}=\mu(x-1)\varphi_t(x).$$

Очевидна властивість твірної функції $\varphi_{t=0}(x)=1$ *стає для цього рівняння початковою умовою. Дана задача Коші має очевидний розв'язок*

$$\varphi_t(x)=\exp(\mu t(x-1)).$$

Оскільки $|x|\le 1$, *то останню функцію можна представити рядом Маклорена за степенями* x

$$\varphi_t(x)=\sum_{k=0}^{\infty}\frac{\mu^k t^k}{k!}\exp(-\mu t)x^k.$$

З іншого боку,

$$\varphi_t(x)=M(x^{\xi_t})=\sum_{k=0}^{\infty}P(\xi_t=k)x^{\xi_t=k}.$$

Отже,

$$P(\xi_t=k)=\frac{\mu^k t^k}{k!}\exp(-\mu t).$$

Розподіл Пуасона має одну рідкісну властивість. Знаючи розподіл випадкового процесу Пуасона ξ_t, можна знайти і закон розподілу довільного набору таких випадкових процесів ξ_{t_1}, ξ_{t_2},

… , ξ_{t_n}. Звичайно, можна здійснити лише обернену процедуру. Знаючи закон розподілу набору випадкових величин, можна знайти закони розподілу лише кожної окремої випадкової величини. Розглянемо це питання детальніше. Нехай маємо систему рівнянь: $\xi_{t_1} = k_1$, $\xi_{t_2} - \xi_{t_1} = k_2 - k_1$, … , $\xi_{t_n} - \xi_{t_{n-1}} = k_n - k_{n-1}$, у правих частинах яких всі величини задані, а у лівих невідомі. Ця система рівнянь абсолютно еквівалентна системі рівнянь $\xi_{t_1} = k_1$, $\xi_{t_2} = k_2$, … , $\xi_{t_n} = k_n$, тобто має ті самі корені. Це означає, що знати набір значень випадкових величин ξ_{t_1}, ξ_{t_2}, … , ξ_{t_n} , це те саме, що знати набір значень випадкових величин ξ_{t_1}, $\xi_{t_2} - \xi_{t_1}$, … , $\xi_{t_n} - \xi_{t_{n-1}}$. У свою чергу це означає, що обидва набори випадкових величини, або випадкових процесів розподілені однаково, тобто

$$P(\xi_{t_1} = k_1, \xi_{t_2} = k_2, ..., \xi_{t_n} = k_n) =$$
$$= P(\xi_{t_1} = k_1, \xi_{t_2} - \xi_{t_1} = k_2 - k_1, ..., \xi_{t_n} - \xi_{t_{n-1}} = k_n - k_{n-1}).$$

Оскільки всі аргументи функції у правій частині є незалежними, то

$$P(\xi_{t_1} = k_1, \xi_{t_2} - \xi_{t_1} = k_2 - k_1, ..., \xi_{t_n} - \xi_{t_{n-1}} = k_n - k_{n-1}) =$$
$$= P(\xi_{t_1} = k_1)\, P(\xi_{t_2} - \xi_{t_1} = k_2 - k_1) ... P(\xi_{t_n} - \xi_{t_{n-1}} = k_n - k_{n-1}).$$

Остаточний результат, очевидно, буде таким

$$P(\xi_{t_1} = k_1, \xi_{t_2} = k_2, ..., \xi_{t_n} = k_n) = \frac{\mu^{k_1} t_1^{k_1}}{k_1!} \exp(-\mu t_1) \times$$
$$\times \frac{\mu^{k_2 - k_1} (t_2 - t_1)^{k_2 - k_{11}}}{(k_2 - k_1)!} \exp(-\mu (t_2 - t_1)) ... \times$$
$$\times \frac{\mu^{k_n - k_{n-1}} (t_n - t_{n-1})^{k_n - k_{n-1}}}{(k_n - k_{n-1})!} \exp(-\mu (t_n - t_{n-1}))$$

$$= P(\xi_{t_1} = k_1, \xi_{t_2} - \xi_{t_1} = k_2 - k_1, ..., \xi_{t_n} - \xi_{t_{n-1}} = k_n - k_{n-1}) .$$

Нагадаємо, що випадкова величина ξ_t для довільного і фіксованого значення t дорівнює кількості стрибків траєкторії процессу ξ_t на відрізку $[0,t)$. Розглянемо тепер набір випадкових величин τ_1, τ_2, … , τ_n, що дорівнюють проміжкам часу між стрибками. Якщо моменти самих стрибків позначити через θ_1, θ_2, … , θ_n (самі ці моменти також є випадковими величинами), то . $\tau_1 = \theta_1$, $\tau_2 = \theta_2 - \theta_1$, … , $\tau_n = \theta_n - \theta_{n-1}$. Без доведення приймемо наступний результат.

Теорема. Випадкові величини τ_1, τ_2, … , τ_n є незалежними і розподіленими за показникових законом з параметром μ.

$$P_{\theta_1,\theta_1,...,\theta_n}(t_1, t_2, ..., t_n) = \mu^n \exp(-\mu t_n) .$$

4.3.Вінерівський процес

Перші три умови існування випадкового процесу Вінера такі ж як і у разі випадкового процесу Пуасона. Четверта ж умова стосується властивостей математичного очікування, а не ймовірності, як у разі процесу Пуасона.

1. ξ_t - випадковий процес з незалежними приростами.
2. ξ_t - випадковий процес, однорідний у часі.
3. $\xi_{t=0}(\omega) = 0$.
4. $M(\xi_h) \approx a\,h$, $M(\xi_h^2) \approx b\,h$, $M(|\,\xi_h\,|^k) \approx 0$, де a і b - довільні дійсні числа, а h - малий проміжок часу, близький до початкового, $k = 3, 4, ...$.

Тут, як і у разі випадкового процесу Пуасона, третя і четверта умови задають поведінку цього процесу у початковий момент часу і у безпосередньому околі цього моменту. Якщо у разі процесу Пуасона цей опис здійснювався мовою ймовірностей відповідних значень випадкової величини, то тепер це здійснюється мовою властивостей математичного очікування. Перша з наближених рівностей пункту чотири свідчить, що у початковий момент часу математичне очікування відповідної випадкової величини (перший початковий момент) дорівнює нулю, і в міру віддалення від цього моменту зростає лінійно, зрозуміло лише при малому віддаленні. Друга наближена рівність означає, що математичне очікування квадрату відповідної випадкової величини (другий початковий момент) також у початковий момент дорівнює нулі і лінійно зростає зі віддаленням від цього початкового моменту, зрозуміло також лише при малому віддаленні. Сукупність перших двох наближених рівностей гарантує, що дисперсія відповідної випадкової величини дорівнює нулю у початковий момент і лінійно зростає в міру віддалення від нього. Дійсно, $D(\xi_h) = M(\xi_h^2) - [M(\xi_h)]^2 \approx b\,h - a^2h^2 \approx b\,h$. Третя наближена рівність означає, що третій і всі старші початкові моменти випадкової величини дорівнюють нулю. Тобто, випадковий процес в околі початкового моменту часу цілком визначається першими двома початковими моментами, або першим початковим і другим центральним Можна сподіватись, що і у наступні моменти часу ця властивість закону розподілу збережеться. Разом з першими двома умовами, ми отримуємо можливість знаходження випадкового процесу Вінера у довільний момент часу. Дійсно, цей процес має очікуваний вигляд

$$F_{\xi_t}(x) = P(\xi_t < x) = \frac{1}{\sqrt{2\pi bt}} \int_{-\infty}^{x} \exp\left(-\frac{(y-at)^2}{2bt} \right) dy,$$

що цілком збігається з нормальною функцією розподілу випадкової величини, якщо вважати at її математичним очікуванням, а bt її дисперсією.

Доведення. *При знаходженні формули для випадкового процесу Пуасона ми використовували твірну функцію дискретної випадкової величини. У разі неперервної випадкової величини аналогічну роль виконує характеристична функція* $f_t(x) = M(\exp(ix\xi_t))$. *Надамо параметру t малий приріст h. Тоді*

$$f_{t+h}(x) = M(\exp(ix\xi_{t+h})) = M(\exp(ix(\xi_{t+h} - \xi_t + \xi_t)) =$$
$$= M(\exp(ix(\xi_{t+h} - \xi_t)\exp(ix(\xi_t)).$$

Оскільки випадковий процес ξ_t є випадковим процес з незалежними приростами, тобто $\xi_{t+h} - \xi_t$ і ξ_t є незалежними величинами, то за властивістю математичного очікування добутку незалежних величин

$$f_{t+h}(x) = M(\exp(ix(\xi_{t+h} - \xi_t)\exp(ix(\xi_t)) =$$
$$= M(\exp(ix(\xi_{t+h} - \xi_t))M(\exp(ix(\xi_t)) = f_t(x)M(\exp(ix(\xi_{t+h} - \xi_t)).$$

Далі скористаємось властивістю однорідності у часі даного випадкового процесу. Це означає, що випадкові величини $\xi_{t+h} - \xi_t$ і $\xi_h - \xi_{t=0}$ однаково розподілені, а отже мають і однакові характеристичні функції $M(\exp(ix(\xi_{t+h} - \xi_t)) = M(\exp(ix(\xi_h - \xi_{t=0}))$. *Далі врахуємо початкову умову для випадкової величини $\xi_{t=0} = 0$. Тоді $M(\exp(ix(\xi_{t+h} - \xi_t)) = M(\exp(ix(\xi_h))$ Нарешті, для малих $h << 1$ останнє математичне очікування можна обчислити безпосередньо. Для цього представимо показникові функцію під знаком математичного очікування рядом Маклорена за* *степенями* x $\exp(ix\xi_h)=\sum_{k=0}^{\infty}\frac{(ix\xi_h)^k}{k!}$. *Математичне очікування такої суми дорівнює сумі математичних очікувань її доданків*

$$M(\exp(ix\xi_h))=M\left(\sum_{k=0}^{\infty}\frac{(ix\xi_h)^k}{k!}\right)=$$

$$=\sum_{k=0}^{\infty}M\left(\frac{(ix\xi_h)^k}{k!}\right)=\sum_{k=0}^{\infty}\frac{(ix)^k}{k!}M(\xi_h^k).$$

Оскільки, за умовою, математичне очікування початкових моментів вище другого поблизу початкового моменту дорівнює нулю, то

$$\sum_{k=0}^{\infty}\frac{(ix)^k}{k!}M(\xi_h^k)\approx 1+ixM(\xi_h)+\frac{(ix)^2}{2!}M(\xi_h^2)=$$

$$=1+ixah-\frac{1}{2}x^2bh,\ |x|\le 1.$$

Таким чином,

$$f_{t+h}(x)\approx f_t(x)\left(1+ixah-\frac{1}{2}x^2bh\right),$$

або

$$\frac{f_{t+h}(x)-f_t(x)}{h}\approx\left(ixa-\frac{1}{2}x^2b\right)f_t(x).$$

Якщо перейти до межі, то остання наближена рівність перетворить у строгу рівність, а різницеве рівняння перетвориться у диференційне

$$\frac{df_t(x)}{dt}=\left(ixa-\frac{1}{2}x^2b\right)f_t(x).$$

Очевидна властивість характеристичної функції $f_{t=0}(x)=1$ *стає для цього рівняння початковою умовою. Дана задача Коші має очевидний розв'язок*

$$f_t(x) = \exp(ixat - \frac{1}{2}x^2bt).$$

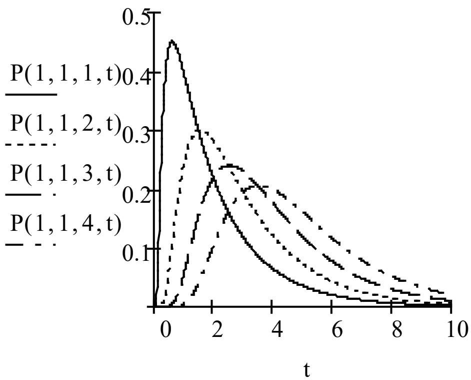
Така характеристична функція якраз і відповідає нормальному розподілу з математичним очікуванням at *і дисперсією* bt. *Отже, густиною цього розподілу є*

$$P_{\xi_t}(x) = \frac{1}{\sqrt{2\pi bt}}\exp\left(-\frac{(x-at)^2}{2bt}\right),$$

а функція розподіл

$$F_{\xi_t}(x) = \frac{1}{\sqrt{2\pi bt}}\int_{-\infty}^{x}\exp\left(-\frac{(y-at)^2}{2bt}\right)dy.$$

Графік густини розподілу випадкового процесу Вінера має вигляд



Мал. 1. *Графік густини розподілу для випадкового процесу Вінера:* $a = b = 1$, $x = 1,2,3,4$.

Бачимо ту саму закономірність, що і у разі розподілу Пуасона: із зростанням проміжку часу зростає і ймовірність більших значень відповідної випадкової величини.

Розподіл випадкового процесу Вінера має таку ж рідкісну властивість, як і розподіл Пуасона. Тут так само, знаючи розподіл випадкового процесу ξ_t, можна знайти і закон розподілу довільного набору таких випадкових процесів ξ_{t_1}, ξ_{t_2}, … , ξ_{t_n}. Дійсно, виходячи з еквівалентності розподілу системи випадкових величин ξ_{t_1}, ξ_{t_2}, … , ξ_{t_n} і їх приростів ξ_{t_1}, $\xi_{t_2}-\xi_{t_1}$, … , $\xi_{t_n}-\xi_{t_{n-1}}$ (дивись попередній параграф), тобто

$$P(\xi_{t_1}<x_1,k_1,\xi_{t_2}<x_2,...,\xi_{t_n}<x_n)=$$
$$=P(\xi_{t_1}<x_1,\xi_{t_2}-\xi_{t_1}<x_2,...,\xi_{t_n}-\xi_{t_{n-1}}<x_n),$$

і незалежності приростів випадкових величин, матимемо

$$P(\xi_{t_1}<x_1,\xi_{t_2}-\xi_{t_1}<x_2,...,\xi_{t_n}-\xi_{t_{n-1}}<x_n)=$$
$$=P(\xi_{t_1}<x_1)\,P(\xi_{t_2}-\xi_{t_1}<x_2)...P(\xi_{t_n}-\xi_{t_{n-1}}<x_n).$$

Звідси з очевидністю випливає, що

$$F_{\xi_{t_1},\xi_{t_2},...,\xi_{t_n}}(x_1,x_2,...,x_n)=\int_G dy_1\int dy_2...\int dy_n\times$$
$$\times\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi b(t_k-t_{k-1})}}\exp\left(-\frac{(y_k-a(t_k-t_{k-1}))^2}{2b(t_k-t_{k-1})}\right).$$

Тут область інтегрування

$$G=\{(y_1,y_2,...y_n):y_1<x_1,y_1+y_2<x_2,...,y_1+y_2+...+y_n<x_n\}.$$

4.4.Рівняння Смолуховського

Як ми вже відмічали, випадкові процеси Пуасона і Вінера мають ту особливість, що знання функції розподілу випадкової величини ξ_t гарантує знання функції розподілу довільного набору випадкових величин ξ_{t_m}. Типовою є ситуація, коли така властивість відсутня. В загальному випадку потрібно задавати ієрархію функцій розподілу все більшої кількості величин ξ_{t_m} і лише у межі $n \to \infty$ вони дозволяють як завгодно точно описати випадковий процес ξ_t. Будемо називати функцію розподілу $F_{\xi_t}(x)$ одно точковою, а функцію розподілу $F_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(x_1, x_2, \dots, x_n)$ n-точковою. Одноточкова функція розподілу задає ймовірність знайти значення випадкової функції в момент часу t такі, що $\xi_t \le x$. Багатоточкова функція розподілу задає ймовірність знайти значення випадкового процесу в момент часу t_1 такі, що $\xi_{t_1} \le x_1$, в момент часу t_2 такі, що $\xi_{t_2} \le x_2$, … , в момент часу t_n такі, що $\xi_{t_n} \le x_n$. Багатоточкова функція розподілу пов'язана з густиною розподілу очевидним співвідношенням

$$F_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(x_1, x_2, \dots, x_n) = \int\limits_{-\infty}^{x_1} dy_1 \int\limits_{-\infty}^{x_2} dy_2 \dots \int\limits_{-\infty}^{x_n} dy_n P_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(y_1, y_2, \dots, y_n).$$

За своїм змістом

$$dp_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(x_1, x_2, \dots, x_n) = P_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

є ймовірністю того, що в момент часу t_1.значення випадкової величини ξ_{t_1} лежать в інтервалі $(x_1, x_1 + dx_1)$, в момент часу t_2 .значення випадкової величини ξ_{t_2} лежать в інтервалі $(x_2, x_2 + dx_2)$, … , в момент часу t_n .значення випадкової величини ξ_{t_n} лежать в інтервалі $(x_n, x_n + dx_n)$.

Ієрархія функцій розподілу має наступні властивості:

1.Густина розподілу завжди невід'ємна

$$P_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(x_1, x_2, \dots, x_n) \ge 0.$$

2.Густина розподілу симетрична відносно перестановки будь-якої пари свої аргументів, наприклад,

$$P_{\xi_{t_1},\xi_{t_2},\dots,\xi_{t_n}}(x_1, x_2, \dots, x_n) = P_{\xi_{t_2},\xi_{t_1},\dots,\xi_{t_n}}(x_2, x_1, \dots, x_n).$$

3.Густина розподілу задовольняє умові узгодженості

$$\int_{-\infty}^{\infty} P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_n}}(x_1,x_2,...,y_n)\,dy_n = P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_{n-1}}}(x_1,x_2,...,x_{n-1}).$$

4.Одноточкова густина розподілу нормована

$$\int_{-\infty}^{x} P_{\xi_t}(y)\,dy = 1.$$

5.Якщо деякі часи збігаються, наприклад $t_1 = t_2$, то

$$P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_n}}(x_1,x_2,...,x_n) = \delta(x_1 - x_2) P_{\xi_{t_2},...,\xi_{t_n}}(x_2,...,x_n).$$

Можна ввести і поняття багатоточкової умовної густини ймовірності. Для цього набір випадкових величин $\xi_{t_1},\xi_{t_2},...,\xi_{t_n}$ слід розділити на дві групи. Перша з них $\xi_{t_1},\xi_{t_2},...,\xi_{t_k}$ задовольняє умові $\xi_{t_1} \le x_1, \xi_{t_2} \le x_2,...,\xi_{t_k} \le x_k$. Для другої групи $\xi_{t_{k+1}},\xi_{t_{k+2}},...,\xi_{t_{k+l}}$ ми шукатимемо густину розподілу ймовірностей при умові, що перша група випадкових величин задовольняє зазначеним умовам. Ця густина розподілу і називатиметься умовною багато точковою густиною ймовірностей. Між умовною густиною розподілу ймовірностей і безумовними густинами існує, за означенням, наступний зв'язок

$$P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_{k+l}}}(x_1,x_2,...,x_k) = P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_k}}(x_1,x_2,...,x_k)\times$$
$$\times P_{\xi_{t_{k+1}},\xi_{t_{k+2}},...,\xi_{t_{k+l}}|\xi_{t_1},\xi_{t_2},...,\xi_{t_k}}(x_{k+1},x_{k+2},...,x_{k+l} \mid x_1,x_2,...,x_k).$$

Умова узгодженості для умовної густини ймовірностей має вигляд

$$\int_{-\infty}^{\infty} P_{\xi_{t_{k+1}},\xi_{t_{k+2}},...,\xi_{t_{k+l}}|\xi_{t_1},\xi_{t_2},...,\xi_{t_k}}(x_{k+1},x_{k+2},...,y_{k+l} \mid x_1,x_2,...,x_k)dy_l =$$
$$= P_{\xi_{t_{k+1}},\xi_{t_{k+2}},...,\xi_{t_{k+l-1}}|\xi_{t_1},\xi_{t_2},...,\xi_{t_k}}(x_{k+1},x_{k+2},...,x_{k+l-1} \mid x_1,x_2,...,x_k).$$

Двоточкова умовна густина ймовірностей ще називається функцією переходу

$$P_{\xi_{t_i},\xi_{t_j}}(x_i,x_j) = P_{\xi_{t_i}}(x_i) P_{\xi_{t_i}|\xi_{t_j}}(x_i \mid x_j).$$

Умова узгодженості для неї, або умова нормування має вигляд

$$\int_{-\infty}^{\infty} P_{\xi_{t_i}|\xi_{t_j}}(x_i \mid x_j)dx_i = 1 .$$

У разі, якщо аргументи функції переходу збігаються, то

$$\lim_{x_j \to x_i + 0} P_{\xi_{t_i}|\xi_{t_j}}(x_i \mid x_j) = \delta(x_i - x_j) .$$

Для стаціонарних випадкових процесів, або як ми раніше їх називали, однорідних у часі

$$P_{\xi_{t_1+\tau},\xi_{t_2+\tau},...,\xi_{t_n+\tau}}(x_1, x_2, ..., x_n) = P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_n}}(x_1, x_2, ..., x_n) .$$

Якщо випадковий процес задовольняє умові

$$P_{\xi_{t_{k+1}}}(x_{k+1} \mid x_1, x_2, ..., x_k) = P_{\xi_{t_{k+1}}}(x_{k+1} \mid x_k) ,$$

то він називається Марківським. Очевидно такий процес в момент часу t_{k+1} залежить лише від його значення в момент часу t_k. Іншими словами, майбутнє залежить лише від теперішнього часу, але не від минулого.

Марківський процесс є складнішим за процесс Пуасона і Вінерівський процесс, оскільки він визначається двома одноточковими густинами розподілу: умовною і безумовною: $P_{\xi_{t_{k+1}}}(x_{k+1} \mid x_k)$ і $P_{\xi_{t_k}}(x_k)$. Двоточкова функція розподілу визначається через них так

$$P_{\xi_{t_1},\xi_{t_2}}(x_1,\ x_2) = P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1)P_{\xi_{t_1}}(x_1) .$$

Триточкова визначиться наступним чином

$$P_{\xi_{t_1},\xi_{t_2},\xi_{t_3}}(x_1,\ x_2, x_3) = P_{\xi_{t_3}|\xi_{t_2}}(x_3 \mid x_2)P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1)P_{\xi_{t_1}}(x_1) .$$

....

Нарешті,

$$P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_n}}(x_1,\ x_2, ..., x_n) = P_{\xi_{t_n}|\xi_{t_{n-1}}}(x_n \mid x_{n-1})...P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1)P_{\xi_{t_1}}(x_1) .$$

Так само, як і випадкові величини, випадкові процеси можна описувати за допомогою числових характеристик. Такими популярними величинами є кореляційні функції, аналоги моментів випадкових величин,

$$K(t_1,\, t_2,...,t_n) = \left\langle \xi_{t_1} \xi_{t_2} ... \xi_{t_n} \right\rangle .$$

Сенс введеного позначення розкривається наступним чином

$$K(t_1,\, t_2,...,t_n) =$$

$$= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 ... \int_{-\infty}^{\infty} dx_n x_1\, x_2\, ...\, x_n P_{\xi_{t_1},\xi_{t_2},...,\xi_{t_n}}(x_1, x_2,..., x_n) .$$

Для отримання рівняння Смолуховського, заінтегруємо вираз для триточкової функції розподілу ймовірності за середнім аргументом

$$\int_{-\infty}^{\infty} P_{\xi_{t_1},\xi_{t_2},\xi_{t_3}}(x_1,\, x_2, x_3) dx_2 = P_{\xi_{t_1}}(x_1) \int_{-\infty}^{\infty} P_{\xi_{t_3}|\xi_{t_2}}(x_3 \mid x_2) P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1) dx_2 .$$

Далі розділимо обидві частини цього співвідношення на $P_{\xi_{t_1}}(x_1)$

$$\frac{P_{\xi_{t_1},\xi_{t_3}}(x_1,\, x_3)}{P_{\xi_{t_1}}(x_1)} = \int_{-\infty}^{\infty} P_{\xi_{t_3}|\xi_{t_2}}(x_3 \mid x_2) P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1) dx_2 .$$

У лівій частині цього рівняння у нас містить функція переходу. Тому остаточний вигляд рівняння Смолуховського для довільного Марківського процесу такий

$$P_{\xi_{t_3}|\xi_{t_1}}(x_3 \mid x_1) = \int_{-\infty}^{\infty} P_{\xi_{t_3}|\xi_{t_2}}(x_3 \mid x_2) P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1) dx_2 .$$

Якщо випадкова величина набуває дискретного набору значень, то рівняння Смолуховського має вигляд

$$P_{\xi_{t_3}|\xi_{t_1}}(x_3 \mid x_1) = \sum_{x_2} P_{\xi_{t_3}|\xi_{t_2}}(x_3 \mid x_2) P_{\xi_{t_2}|\xi_{t_1}}(x_2 \mid x_1) .$$

4.5.Броунівський рух

Одновимірний Броунівський рух описується рівнянням Ньютона, де невідомою функцією є координата частинки. Якщо це рівняння записати відносно швидкості, то воно називається рівнянням Ланжевена і має вигляд

$$\frac{d\xi_t}{dt} + \nu\xi_t = f(t),$$

$$\xi_{t=t_0} = x_0 .$$

Тут ξ_t - швидкість Броунівської частинки, ν - числовий параметр, що визначає макроскопічну силу, яка діє на Броунівську частинку з боку рідини. Це сила опору рідини руху частинки. $f(t)$ - випадкова сила, що діє на броунівську частинку, у висліді зіткнення з нею окремих молекул рідини. Оскільки рівняння лінійне, то його загальний розв'язок є сумою загального розв'язку однорідного рівняння і частинного розв'язку неоднорідного. Загальний розв'язок однорідного рівняння, очевидно, такий

$$\xi_t = A\exp(-\nu t).$$

Тут A – довільна стала. Частинний розв'язок неоднорідного знайдемо методом варіації довільної сталої. Нехай $A = A(t)$. Тоді шуканий частинний розв'язок запишемо так

$$\xi_t = A(t)\exp(-\nu t).$$

Після підстановки його у рівняння матимемо

$$\frac{dA(t)}{dt}\exp(-\nu t) = f(t).$$

Інтегруючи це рівняння отримаємо

$$A(t) = \int_B^t \exp(\nu\tau) f(\tau) d\tau .$$

Сам частинний розв'язок тепер можна записати так

$$\xi_t = \int_B^t \exp(-\nu(t-\tau)) f(\tau) d\tau .$$

Тут B – довільна стала. Оскільки мова йде про довільний частинний розв'язок, то цю довільну сталу можна обрати довільним чином, наприклад, $B = t_0$. Загальний розв'язок матиме вигляд

$$\xi_t = A\exp(-\nu t) + \int_{t_0}^{t} \exp(-\nu(t-\tau)) f(\tau) d\tau .$$

Підстановка загального розв'язку у початкову умову дає шуканий частинний розв'язок вихідної задачі Коші

$$\xi_t = x_0 \exp(-\nu(t-t_0)) + \int_{t_0}^{t} \exp(-\nu(t-\tau)) f(\tau) d\tau .$$

Якщо сила $f(\tau) = \overline{f(\tau)}$ є невипадковою функцією, а цілком визначеною, то невипадковою функцією буде і функція $\xi_t = \overline{\xi_t}$. У цьому разі мова йтитиме не про ймовірність того, що $\xi_t \le x$, а про те, що $\overline{\xi_t} = x$. Ймовірність останньої рівності – єдина, відмінна від нуля. Всі інші ймовірності дорівнюють нулю. Так само має бути $\overline{\xi_{t_0}} = x_0$. Оскільки функція переходу є нормованою, то єдиним шляхом задовольнити зазначені умови, є функція переходу у вигляді дельта-функції Дірака

$$\overline{P_{\xi_t|\xi_{t_0}}}(x \mid x_0) = \delta(x - \overline{\xi_t}) .$$

Очевидно, що при $t = t_0$ тематично виконуватиметься умова $\overline{\xi_{t_0}} = x_0$. Випадок стохастичної сили можна отримати узагальненням виразу для нестохастичної сили. Кожний варіант стохастичної сили породжуватиме функцію переходу з дельта-функцією, що обирає з усіх можливих траєкторій таку, що відповідає даній реалізації стохастичної сили. Для урахування всіх реалізацій стохастичної сили і породжених нею траєкторій, потрібно лише усереднити отриманий вище вираз за всіма реалізаціями стохастичної сили

$$P_{\xi_t|\xi_{t_0}}(x \mid x_0) = \left\langle \delta(x - \xi_t) \right\rangle .$$

Для виконання заявленого усереднення використаємо представлення дельта-функції інтегралом Фур'є

$$\delta(x-a)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\exp(i(x-a)k)dk\,.$$

Тоді для функції переходу дістанемо наступний вираз

$$P_{\xi_t|\xi_{t_0}}(x\,|\,x_0)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\exp(i(x-x_0\exp(-\nu(t-t_0)))k)dk\times$$

$$\times\left\langle\exp\left(ik\int\limits_{t_0}^{t}\exp(-\nu(t-\tau))f(\tau)d\tau\right)\right\rangle.$$

Розглянемо найпростіший варіант стохастичної функції. Це так званий білий шум – максимально можливий випадковий характер функції. Білий шум найкраще задається своїми числовими характеристиками – кореляційними функціями. Всі непарні кореляційні функції, починаючи з першої, дорівнюють нулю

$$\left\langle f(t_1)f(t_2)...f(t_{2n+1})\right\rangle=0\,.$$

Всі парні кореляційні функції відмінні від нуля і цілком визначаються через двоточкову

$$\left\langle f(t_1)f(t_2)\right\rangle=\sigma_f\delta(t_1-t_2)\,,$$

$$\left\langle f(t_1)f(t_2)...f(t_{2n-1})f(t_{2n})\right\rangle=\sigma_f^n\delta(t_1-t_2)\delta(t_3-t_4)...\delta(t_{2n-1}-t_{2n})\,.$$

Експоненту під знаком середнього представимо рядом Маклорена

$$\left\langle\exp\left(ik\int\limits_{t_0}^{t}\exp(-\nu(t-\tau))f(\tau)d\tau\right)\right\rangle=$$

$$=\left\langle\sum_{m=0}^{\infty}\frac{1}{m!}\left[ik\int\limits_{t_0}^{t}\exp(-\nu(t-\tau))f(\tau)d\tau\right]^m\right\rangle=$$

$$=\sum_{m=0}^{\infty}\frac{1}{m!}\left[ik\left\langle\int\limits_{t_0}^{t}\exp(-\nu(t-\tau))f(\tau)d\tau\right\rangle\right]^m=$$

$$= \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \left\langle f(\tau_1) f(\tau_2) \ldots f(\tau_m) \right\rangle \times$$

$$\times \int_{t_0}^{t} d\tau_1 \exp(-\nu(t-\tau_1)) \int_{t_0}^{t} d\tau_2 \exp(-\nu(t-\tau_2)) \ldots \int_{t_0}^{t} d\tau_m \exp(-\nu(t-\tau_m)) .$$

Наявність дельта-функцій під знаком інтегралів дозволяє їх легко обчислити і остаточний результат буде таким

$$P_{\xi_t|\xi_{t_0}}(x \mid x_0) = \frac{1}{\sqrt{\pi\sigma\nu^{-1}(1-\exp(-2\nu(t-t_0)))}} \times$$

$$\times \exp\left(-\frac{1}{\nu^{-1}\sigma} \frac{(x - x_0 \exp(-\nu(t-t_0)))}{1-\exp(-2\nu(t-t_0))} \right).$$

Випадковий процес, що описується даною формулою називається випадковим процесом Орнштейна-Уленбека. Якщо опір руху Броунівських частинок відсутній $\nu = 0$, то ми отримуємо випадковий процес Вінера.

Додатки

1.Розподіл неприємностей

У своєму житті людина зустрічається з неприємними подіями. У зв'язку з цим знайдемо густину розподілу ймовірності життя людини без неприємних подій. В якості простору елементарних подій візьмемо додатну піввісь t. Нехай $\overline{[0,t]}$ - подія, що полягає у тому, що за проміжок часу $[0,t]$ відбулась неприємна подія. Подібною до неї є і подія $\overline{(t,t+dt]}$ та $\overline{[0,t+dt]}$.

Очевидно

$$\overline{[0,t+dt]} = \overline{[0,t]} \cup \overline{[t,t+dt]}$$

і події $\overline{[0,t]}$, і $\overline{[t,t+dt]}$ є незмінним. Відповідно $\overline{[0,t]}$, $\overline{(t,t+dt]}$, $\overline{[0,t+dt]}$ - події, які полягають у тому, що за відповідні проміжки часу неприємні події не відбудуться. Очевидно

$$\overline{[0,t+dt]} = \overline{[0,t]} \cap \overline{(t,t+dt]}$$

і події $\overline{[0,t]}$, $\overline{(t,t+dt]}$ є незалежними.

Нехай тепер $w(t) = P(0 \le \xi \le t)$ є ймовірністю того, що за проміжок часу $[0,t]$ неприємна подія не відбудеться. Відповідно до теореми множення ймовірностей двох незалежних подій.

$$W([0,t+dt]) = W([0,t])W([t,t+dt]),$$

або

$$W(t+dt) = W(t)W(dt).$$

Функції $W(t+dt)$ і $W(dt)$ зручно розвинути в ряд Тейлора за степенями dt, обмежившись при цьому лінійними за dt членами. Отже

$$W(dt) = W(0) + W'(0)dt,$$

або

$$W(dt) = 1 - adt\,.$$

В останні рівності було враховано, що за нульовий проміжок часу неприємна подія не відбудеться і мова йде про достовірну подію. Відповідно чи до однієї з аксіом ймовірності ймовірність достовірної події є одиниця $W(0) = 1$. Оскільки за іншою аксіомою ймовірності остання не перевищує одиниці $W(dt) \le 1$, то $W'(0) = -a \le 0$ для $a \ge 0$. Аналогічно розвинемо і функцію $W(t + dt)$

$$W(t + dt) = W(t) + W'(t)dt\,.$$

Після підстановки обох розвинень в основне рівняння отримаємо

$$W'(t) = -aW(t)\,.$$

Це диференційне рівняння має наступний загальний розв'язок

$$W(t) = A\exp(-at)\,.$$

Сталу A можна знайти із вже обговорюваної вище умови $W(0) = 1$, отже $A = 1$. Таким чином ймовірність відсутності неприємних подій описується наступним законом

$$W(t) = \exp(-at)$$

і швидкість спадання з часом цієї ймовірності визначається сталою a. Визначимо її сенс. Для цього розглянемо функцію

$$F_\xi(t) = 1 - W(t) = 1 - \exp(-at)\,,$$

яка є функцією розподілу ймовірностей для неприємних подій, тобто ймовірністю того, що неприємна подія ξ відбудеться за проміжок часу $[0, t]$. Знайдемо математичне очікування цієї неприємної події, тобто середній проміжок часу за який відбудеться хоча б одна подія. За означенням

$$\tau = M\xi = \int_0^\infty tP_\xi(t)dt$$

де густина ймовірності

$$P_\xi(t) = dF_\xi(t)/dt = a\exp(-at)\,.$$

Отже,

$$\tau = a\int_0^\infty t\exp(-at)dt = \frac{1}{a}.$$

Таким чином,

$$\mathrm{P}_\xi(t) = \tau^{-1}\exp(-t/\tau)$$

і

$$W(t) = \exp(-t/\tau)$$

Зрозуміло, що конкретне значення часу релаксації, або середньому проміжку часу між неприємностями τ залишається невизначеним. Останнє потребує додаткової інформації про перебіг життя конкретної людини.

2.Класичний електронний газ

У теорії електронного газу використовують різні функції розподілу. Розглянемо деякі з них. У невзаємодіючому електронному газі електрон має лише кінетичну енергію, яку можна записати так $\varepsilon_k = \dfrac{\hbar^2 k^2}{2m}$, де хвильовий вектор електрона $\mathbf{k}$ є добрим квантовим числом, що характеризує стан, у якому знаходиться електрон. Разом із спіновим квантовим числом $s = \pm 1/2$, хвильовий вектор утворює повний набір квантових чисел для стану електрона. Оскільки енергія електрона у даному разі не залежить від спінового квантового числа, то останнє ми явно писати не будемо.

Розглянемо функцію $n(k)$, що має сенс середньої кількості електронів у стані з хвильовим вектором $\mathbf{k}$. Нехай загальна кількість електронів N, тоді

$$\sum_{\mathbf{k}} n(k) = N\,.$$

Розглянемо класичний електронний газ. З загальних принципів статистичної фізики випливає, що функція $n(k)$ має вигляд

$$n\left(k\right) = C\exp\left(-\frac{\varepsilon_k}{k_B T}\right)\,.$$

Сталу величину C можна визначити з попередньої формули

$$\sum_{\mathbf{k}} n(\mathbf{k}) = \frac{2V}{(2\pi)^3}\int_0^\infty dk\, k^2 n(k)\int_\Omega d\Omega = N\,,$$

де ми перейшли від підсумовування до інтегрування у сферичній системі координат, V - об'єм системи. Оскільки закон розподілу не залежить від кутів, то інтеграл за тілесним кутом

$$\int_\Omega d\Omega = 4\pi$$

Тоді

$$\sum_{\mathbf{k}} n\left(k\right) = \frac{2V}{\left(2\pi\right)^3}4\pi C\int_0^\infty dk k^2 \exp\left(-\frac{\hbar^2 k^2}{2mk_B T}\right)$$

Для обчислення інтегралу використаємо відомий результат для інтегралу Пуассона

$$\int_0^\infty dx \exp\left(-\alpha x^2\right) = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Здиференціюємо обидві частини цього виразу за α . Матимемо

$$\int_0^\infty dx x^2 \exp\left(-\alpha x^2\right) = \frac{1}{4}\frac{\pi^{1/2}}{\alpha^{3/2}}$$

Якщо покласти $\alpha = \dfrac{\hbar^2}{2mk_B T}$, то

$$\int_0^\infty dk k^2 \exp\left(-\frac{\hbar^2 k^2}{2mk_B T}\right) = \frac{1}{4}\frac{\pi^{1/2}\left(2mk_B T\right)^{3/2}}{\hbar^3}$$

Таким чином

$$\frac{2V}{(2\pi)^3}4\pi C\cdot\frac{1}{4}\frac{\pi^{1/2}\left(2mk_BT\right)^{3/2}}{\hbar^3}=N$$

Звідки

$$C=4\pi^3\hbar^3 n\frac{1}{\left(2\pi mk_BT\right)^{3/2}}$$

Середня густина матиме вигляд

$$n(k)=\frac{4\pi^3\hbar^3 n}{\left(2\pi mk_BT\right)^{3/2}}\exp\left(-\frac{\hbar^2k^2}{2mk_BT}\right)$$

Якщо розглядати електронний газ з квантовомеханічної точки зору, виходячи з існування дискретного набору станів, де знаходиться електрон, то статистичні властивості електронного газу доцільно описувати за допомогою середньої густини $n(k)$. Якщо розглядати електронний газ з класичної точки зору, то у відсутності взаємодії єдиним параметром, що характеризує стан електрона, є його швидкість. Оскільки цей параметр змінюється неперервно, то статистичні властивості електронного газу доцільно описувати за допомогою густини розподілу $f(v)$. За означенням

$$dp(\mathbf{v})=f(\mathbf{v})d\mathbf{v}$$

- ймовірність знаходження електрону у стані з швидкістю $\mathbf{v}$, що належить інтервалу $\left[\mathbf{v},\mathbf{v}+d\mathbf{v}\right]$. З загальних принципів статистичної фізики випливає, що густина розподілу

$$f\left(v\right)=C\exp\left[-\frac{\varepsilon_k(v)}{k_BT}\right],$$

де кінетична енергія

$$\varepsilon(v)=\frac{mv^2}{2}$$

Стала С може бути знайдена з умови нормування

$$\int f(\mathbf{v})d\mathbf{v} = 1$$

або

$$4\pi \int_0^\infty dv v^2 f(v) = 1$$

Абсолютно аналогічно попередньому знаходимо

$$f(v) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp\left(-\frac{mv^2}{2k_B T} \right)$$

Густина цього розподілу у статистичній фізиці називається густиною розподілу Максвела.

Знайдемо деякі властивості електронного газу. Середня швидкість електрона. У Декартовій системі координат середня швидкість електрона визначається так

$$\langle \mathbf{v} \rangle = \mathbf{i}\langle v_x \rangle + \mathbf{j}\langle v_y \rangle + \mathbf{k}\langle v_z \rangle$$

Тут v_x, v_y, v_z - координати вектора швидкості $\mathbf{v}$, $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ - одиничні орти Декартової системи координат. За означенням математичного очікування

$$\langle v_x \rangle = \int v_x f(v) d\mathbf{v} = \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_z v_x f(v)$$

Оскільки за змістом v_x інтегранда є непарною функцією, то відповідний інтеграл у симетричних межах є нулем. Отже

$$\langle v_x \rangle = 0 .$$

Аналогічно

$$\langle v_y \rangle = \langle v_z \rangle = 0$$

Знайдемо тепер середню величину абсолютної швидкості електрона

$$\langle v\rangle = \int v f(v) d\mathbf{v}$$

Тут доцільно використати сферичну систему координат. Для обчислення інтеграла зручно перейти до нової змінної інтегрування $x = v^2$. Тоді

$$\langle v\rangle == 2\pi\left(\frac{m}{2\pi k_B T}\right)^{3/2}\int_0^\infty dxx\exp\left(-\frac{mx}{2k_B T}\right) =$$

$$= 2\pi\left(\frac{m}{2\pi k_B T}\right)^{3/2}\left\{-\frac{2k_B T}{m}x\exp\left(-\frac{mx}{2k_B T}\right)\Bigg|_0^\infty + \frac{2k_B T}{m}\exp\left(-\frac{mx}{2k_B T}\right)\int_0^\infty \exp\left(-\frac{mx}{2k_B T}\right)dx\right\} =$$

$$= 2\pi\left(\frac{m}{2\pi k_B T}\right)^{3/2}\left(\frac{2k_B T}{m}\right)^2 .$$

Остаточно

$$\langle v\rangle = \left(\frac{2k_B T}{m}\right)^{1/2}\frac{2}{\sqrt{\pi}}, \qquad 2/\sqrt{\pi}\approx 1.13 .$$

Знайдемо середню кінетичну енергію електрона. За означенням математичного очікування

$$\langle \varepsilon_k\rangle = \int \frac{mv^2}{2} f(v) d\mathbf{v} = 2\pi m\int_0^\infty dv v^4 f(v) =$$

$$= 2\pi m\left(\frac{m}{2\pi k_B T}\right)^{3/2}\int_0^\infty v^4\exp\left[-\frac{mv^2}{2k_B T}\right]dv$$

Для обчислення останнього інтеграла знову використаємо інтеграл Пуассона

$$\int_0^\infty dx\exp\left(-\alpha x^2\right) = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Здиференціювавши його двічі за параметром отримаємо

$$\frac{d^2}{d\alpha^2}\int_0^\infty dx e^{-\alpha x^2} = \int_0^\infty dx x^4 e^{-\alpha x^2} = \frac{3\pi^{1/2}}{8}\alpha^{-5/2}$$

Отже

$$\left\langle \varepsilon_k \right\rangle = 2\pi m \left(\frac{m}{2\pi k_B T} \right)^{3/2} \frac{3\pi^{1/2}}{8} \left(\frac{2k_B T}{m} \right)^{5/2} .$$

Середня квадратична швидкість

$$\sqrt{\left\langle v^2 \right\rangle} = \sqrt{\frac{2}{m} \varepsilon_{kin}} = \sqrt{\frac{3k_B T}{m}}$$

Очевидно

$$\sqrt{\left\langle v^2 \right\rangle} = \sqrt{\frac{3\pi}{8}} v .$$

Дисперсія швидкості електрона. За означенням дисперсії

$$\left\langle \left[v - \left\langle v \right\rangle \right]^2 \right\rangle = \left\langle v^2 \right\rangle - \left\langle v \right\rangle^2$$

Оскільки раніше нами були отримані наступні результати:

$$\left\langle v \right\rangle = \left(\frac{8k_B T}{\pi m} \right)^{1/2} ,$$

$$\left\langle \varepsilon_k \right\rangle = \frac{m}{2} \left\langle v^2 \right\rangle = \frac{3}{2} k_B T ,$$

то

$$\left\langle v^2 \right\rangle = \frac{3k_B T}{m}$$

$$\left\langle \left[v - \left\langle v \right\rangle \right]^2 \right\rangle = \frac{3k_B T}{m} - \frac{8k_B T}{m\pi} = \frac{k_B T}{m} \left(3 - \frac{8}{\pi} \right)$$

$$\left\langle \left[v - \left\langle v \right\rangle \right]^2 \right\rangle = \frac{k_B T}{m} \left(3 - \frac{8}{\pi} \right)$$

Дисперсія кінетичної енергії електрона.

$$\left\langle \left[\varepsilon_{kin} - \left\langle \varepsilon_{kin} \right\rangle \right]^2 \right\rangle = \left\langle \varepsilon_{kin}{}^2 \right\rangle - \left\langle \varepsilon_{kin} \right\rangle^2 =$$

$$= \frac{m^2}{4} \left[\left\langle v^4 \right\rangle - \left\langle v^2 \right\rangle^2 \right] .$$

Для середнього квадрату швидкості раніше нами був отриманий результат

$$\left\langle v^2 \right\rangle = \frac{3k_B T}{m} \ .$$

Обчислимо $\left\langle v^4 \right\rangle$. За означенням

$$\left\langle v^4 \right\rangle = \int v^4 f(v) d\mathbf{v} = 4\pi \int_0^\infty dv v^6 f(v) =$$

$$= 4\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \int_0^\infty dv v^6 \exp\left[-\frac{mv^2}{2k_B T} \right]$$

Для обчислення останнього інтегралу знову використаємо інтеграл Пуассона

$$\int_0^\infty dx \exp(-\alpha x^2) = \frac{\pi^{1/2}}{2} \alpha^{-1/2} .$$

Здиференціювавши його тричі за α отримаємо

$$-\frac{d^3}{d\alpha^3} \int_0^\infty dx \exp(-\alpha x^2) = \int_0^\infty dx x^6 \exp(-\alpha x^2) = \frac{15\pi^{1/2}}{16} \alpha^{-7/2}$$

Отже, поклавши $\alpha = \frac{m}{2k_B T}$ маємо

$$\left\langle v^4 \right\rangle = 4\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \frac{15\pi^{1/2}}{16} \left(\frac{2k_B T}{m} \right)^{7/2} =$$

$$= \frac{15\pi^{1/2}}{4} \left(\frac{2k_B T}{m} \right)^2 = 15 \left(\frac{k_B T}{m} \right)^2$$

Остаточно

$$\left\langle \left[\varepsilon_{kin} - \left\langle \varepsilon_{kin} \right\rangle \right]^2 \right\rangle = \frac{m^2}{4} \left\{ 15 \left(\frac{k_B T}{m} \right)^2 - 9 \left(\frac{k_B T}{m} \right)^2 \right\} = \frac{3}{2} \left(k_B T \right)^2$$

Знайдемо рівняння стану ідеального газу. Тиск газу – це імпульс, переданий молекулами газу за одиницю часу ділянці стіни одиничної площини. Нехай стіна розташована перпендикулярно осі x. З розподілу Максвелла, що описує розподіл всіх компонент швидкості молекул, перейдемо до розподілу Максвелла, що описує лише розподіл v_x складової швидкості молекул. Для цього обчислимо наступний інтеграл

$$f(v_X) = \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_z f(v_x, v_y, v_z)$$

Очевидно

$$f(v_x) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} \exp\left(-\frac{mv_x^2}{2k_B T}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{mv_y^2}{2k_B T}\right) dv_y \int_{-\infty}^{\infty} \exp\left(-\frac{mv_z^2}{2k_B T}\right) dv_z$$

Кожний з цих інтегралів є інтегралом Пуассона. Отже,

$$f(v_x) = \left(\frac{m}{2\pi k_B T}\right)^{1/2} \exp\left(-\frac{mv_x^2}{2k_B T}\right)$$

Очевидно, що отриманий нами розподіл є нормованим

$$\int_{-\infty}^{\infty} f(v_x) dv_x = 1$$

Помноживши обидві частини рівності на середню густину молекул отримаємо

$$n = \int_{-\infty}^{\infty} dn(v_x),$$

де

$$dn(v_x) = nf(v_x) dv_x$$

- кількість молекул одиниці об'єму, компонента швидкості v_x яких в інтервалі $\left[v_x, v_x + dv_x\right]$. Кількість молекул з швидкістю v_x, що за одиницю часу досягти стінки і передати свій імпульс одиничній площадці, очевидно визначається об'ємом паралелепіпеда висотою v_x і одиничною площиною основи, що лежить на стінці $V = v_x \cdot 1$. Тоді $v_x dn(v_x)$ очевидно і буде визначати ту кількість молекул з швидкістю v_x, що зіткнеться з одиничною площадкою стінки за 1 секунду. Кожна молекула передасть при зіткненні імпульс $2mv_x$. Тиск же визначиться наступним інтегралом

$$P = n\int_{-\infty}^{\infty} dv_x 2mv_x \cdot v_x f(v_x)$$

або

$$P = 2m\left(\frac{m}{2\pi k_B T}\right)^{1/2} n\int_{-\infty}^{\infty} v_x^{\ 2} \exp\left(-\frac{mv_x^2}{2k_B T}\right)dv_x$$

Оскільки

$$\int_0^{\infty} dx x^2 \exp\left(-\alpha x^2\right) = \frac{\pi^{1/2}}{4}\alpha^{-3/2},$$

то

$$P = 2m\left(\frac{m}{2\pi k_B T}\right)^{1/2} \frac{\pi^{1/2}}{4} n\left(\frac{2k_B T}{m}\right)^{3/2}$$

Після спрощення

$$P = nk_B T$$

3.Інформаційна ентропія і свобода вибору

Статистична фізика впевнено завойовує нові позиції в царинах, що здавалося б не мають до неї жодного стосунку. Одним з засадничих понять статистичної фізики є поняття ентропії. У 1877 році геніальний австрійський фізик Людвіг Едуард Больцман першим зрозумів зв’язок ентропії фізичної системи з ймовірністю її перебування в тому, чи іншому макроскопічному стані, пов’язаною з кількістю мікростанів, які реалізують даний макроскопічний стан. У 1948 році знаменитий американський електротехнік і математик Клод Елвуд Шеннон запропонував використовувати поняття ентропії для оцінки невизначеності інформації про ту, чи іншу подію. Тим самим він започаткував нову математичну дисципліну – теорію інформації, де ентропія отримала назву інформаційної ентропії. Зв'язок інформаційної ентропії з ймовірністю настання тої, чи іншої події він запропонував у такому ж вигляді як і зв'язок перебування системи в тому, чи іншому макроскопічному стані, тобто, фактично використав для інформаційної ентропії формулу Больцмана. Розглянемо формулу для інформаційної ентропії.

Інформаційна ентропія – ентропія вибору

Нехай ми проводимо випадковий експеримент з наслідками E_1, E_2, …, E_n, що можуть реалізовуватися з ймовірностями p_1, p_2, …, p_n. Тоді інформація, отримана нами у висліді цього експерименту є випадковою величиною, що приймає значення $I(E_i)$ при настанні наслідку експерименту E_i. При цьому

$$I(E_i) = -\log_2(p_i) .$$

Математичне очікування цієї інформації (інформаційна ентропія), тобто середня кількість інформації, що припадає на один наслідок експерименту, визначається стандартним чином

$$M(I(E_i)) = \sum_{i=1}^{n} p_i I(E_i)$$

Останній результат можна записати і у формі, прийнятій для теорії інформації

$$S = -\sum_{i=1}^{m} p_i \log_2(p_i) .$$

Багато політичних процесів нагадують випадковий експеримент. Зокрема, на нашу думку, таким процесом є вибори різних рівнів. Якщо можливими наслідками цього політичного процесу вважати перемогу того, чи іншого кандидату, тобто події E_1, E_2, …, E_n з ймовірностями p_1, p_2, …, p_n, то для оцінки результатів виборів можна використати попередню формулу. Тепер було б доцільно використовувати замість терміну інформаційна ентропія терміну ентропія виборів, або індекс свободи виборів. Ентропія виборів залежить від двох основних факторів: кількості кандидатів і ймовірностей обрання того, чи іншого кандидата. Ентропія виборів при фіксованій кількості кандидатів досягає максимуму, якщо голоси між різними кандидатами розподілилися порівно, тобто $p_1 = p_2 = ... = p_n$. Цю універсальну властивість ентропії застеріг ще Людвіг Едуард Больцман. У цьому разі формула для ентропії виборів матиме вигляд

$$S = \log_2(n).$$

Тобто із зростанням кількості кандидатів ентропія виборів зростає, як логарифм їх кількості. Це є ніщо інше, як знаменита формула Ральфа Гартлі.

Таким чином, ентропія виборів характеризує рівень невизначеності результатів виборів. Чим більше кандидатів на вибірну посаду і чим рівномірніше розподілені ймовірності перемоги різних кандидатів, тим більшою є ентропія виборів, або індекс свободи виборів. У разі двох кандидатів, що мають рівні шанси на перемогу, тобто у найпростішому політичному випадковому експерименті для ентропії виборів, так само, як і для інформаційної ентропії, ми отримуємо результат 1. В теорії інформації така кількість інформації називається бітом. У нашому випадку її доцільніше було б назвати інакше, наприклад, фрід – від перших літер англійського слова freedom – свобода. Саме для того, щоб у найпростішому політичному експериментів ми отримали відповідь одиницю, логарифм у формулі для ентропії виборів доцільно брати за основою два.

Принагідно зауважимо, що у деяких країнах, рівень демократії у яких вважається високим, фактично вибори і відбуваються між двома кандидатами з приблизно рівними шансами на перемогу. Ясно, що у цьому разі ентропія вибору буде доволі малою, порівняно наприклад, з Україною.

У Радянському Союзі вибори завжди проводились на безальтернативній основі. Це означає, що одна з ймовірностей, наприклад, $p_1 = 1$, а всі інші дорівнюють нулю. У такому разі формула для ентропії виборів дає нульовий результат.

Якщо вибори проходять у два тури, то для другого туру слід використати ту саму формулу для ентропії виборів, а результати, відповідно до універсальної властивості ентропії двох незалежних підсистем однієї системи, скласти

$$S = S_1 + S_2,$$

де

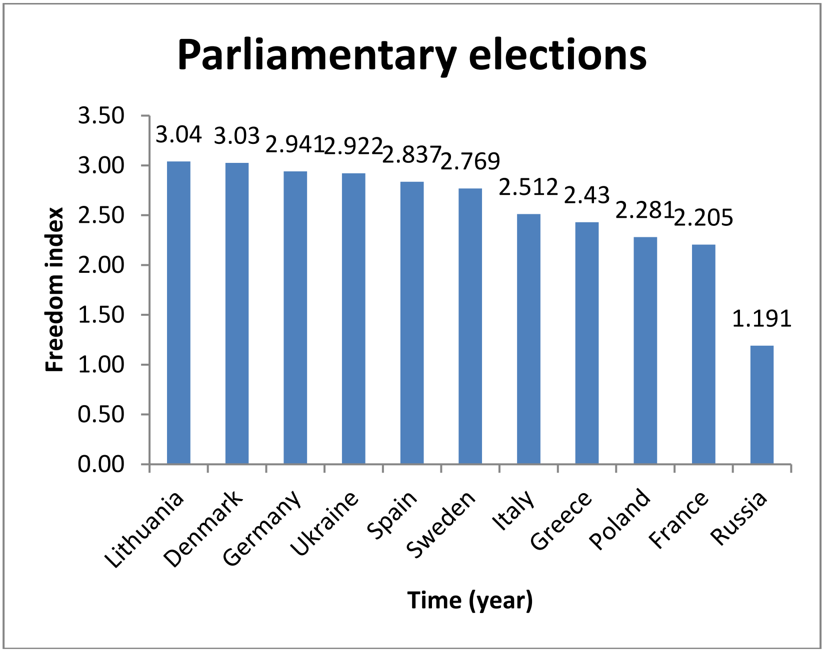
$$S_1 = -\sum_{i=1}^{m} p_i \log_2(p_i),$$

$$S_2 = -\sum_{i=1}^{2} P_i \log_2(P_i).$$

З нашої точки зори ступінь невизначеності результатів виборів свідчить про рівень свободи виборця на таких виборах, тобто про рівень свободи самих виборів. Тому, оскільки конкуруючих варіантів означення з очевидних причин немає, ми пропонуємо ентропію виборів називати індексом свободи виборів. При цьому, якщо виборець голосує за i–го кандидата, то він реалізує свою свободу вибору у кількісному вимірі як $-\log_2(p_i)$.

Вибори в Україні, Росії і країнах Євросоюзу

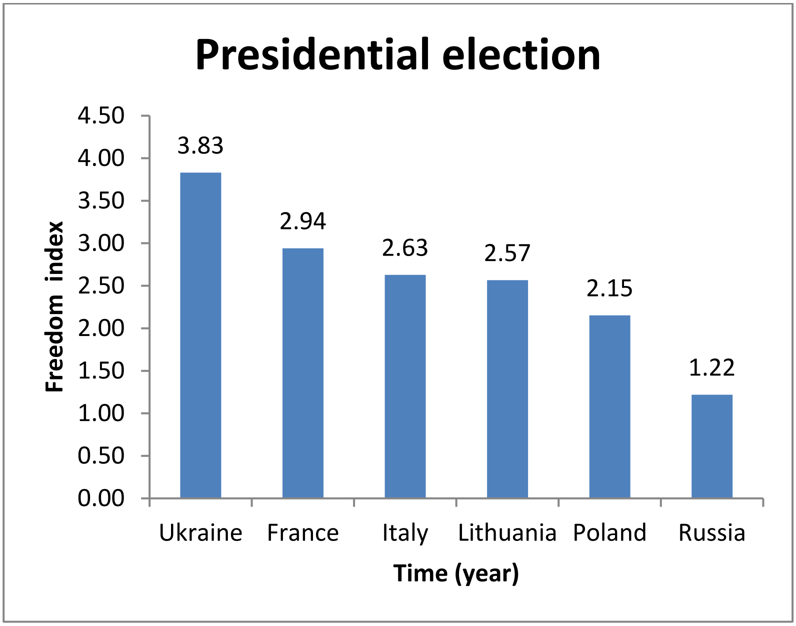
Застосуємо запропоновані нами формули для оцінки індексу свободи виборів у провідних країнах Євросоюзу, а також в Україні і Росії. Для простоти візьмемо лише останні вибори у країнах Європи. Для парламентських виборів матимемо наступний результат.



Мал. 1.

З наведеної діаграми видно, що для всіх розглянутих країн, а це провідні країни Євросоюзу, характерні високі значення індексу свободи виборів. Різниця між найбільшим і найменшим значеннями цього індексу порівняно невелика. Приємно, що лише три країни Євросоюзу мають значення індексу свободи вищі за його значення для України, а решта шість – нижчі. Росія ж у цьому переліку країн посідає останнє місце з великим відривом від країни Євросоюзу з найменшим значенням індексу свободи виборів – Франції. Причому, індекси свободи виборів для Франції і Росії відрізняються майже вдвічі на користь Франції.

Не в усіх країнах Євросоюзу президенти обираються шляхом прямих і загальних виборів. Серед вже розглянутих країн Євросоюзу таких країн є лише чотири. Відповідна діаграма щодо президентських виборів має наступний вигляд.



Мал. 2.

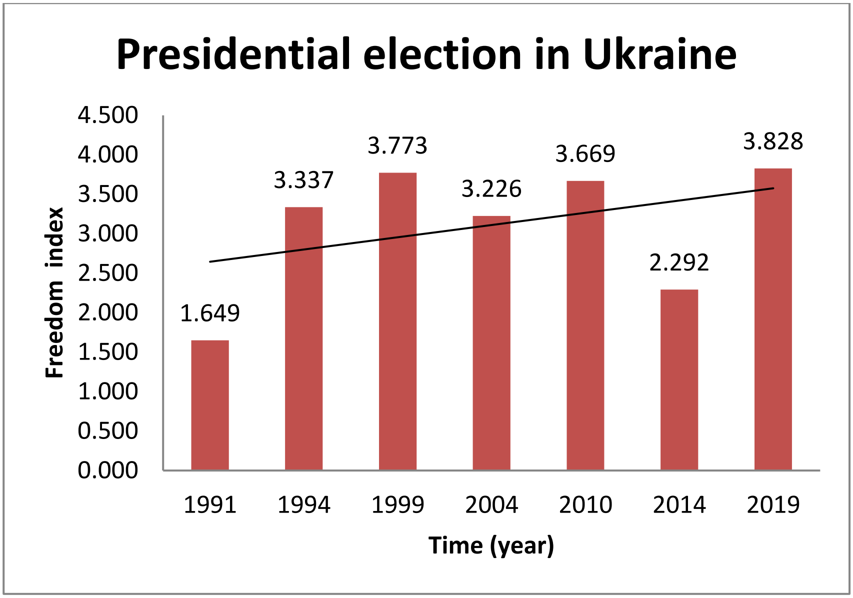
Тут уже безперечним лідером є Україна, причому з великим відривом від решти країн. Інші країни Євросоюзу знову демонструють невелику розкиданість відповідних значень індексу свободи виборів, рівень якого в цілому є доволі високим. На останньому місті за значенням індексу свободи виборів знову опинилась Росія. І знову індекси свободи виборів для країни Євросоюзу з найменшим індексом свободи виборів – Польщі і Росії відрізняються майже вдвічі тепер уже на користь Польщі.

Зауважимо, що необхідні для побудови діаграм данні легко доступні з інтернету і ми не робили на них спеціальних посилань.

Для того, щоб зрозуміти, чи величина індексу свободи виборів відображає менталітет народів і, чи може вона суттєво змінюватись у часі, потрібно проаналізувати поведінку індексу за тривалий проміжок часу в одній і тій самій країні, або у декількох країнах із спільним минулим. Для такого порівняння ми обрали Україну і Росію. Вони сотні років входили до складу однієї імперії, їх мови належать до однієї мовної групи – слав'янської, в кожній з цих країн ще присутній значний відсоток населення, що ідеологічно сформувалось в однакових політичних умовах.

Президентські і парламентські вибори в Україні

На наступній діаграмі наведені результати президентських виборів в Україні за роки незалежності.



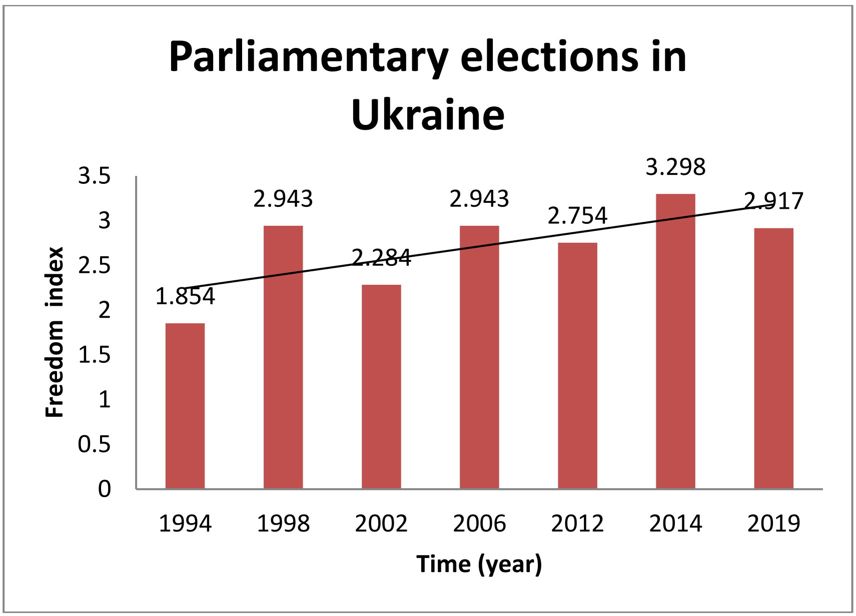
Мал. 3.

З діаграми видно, що індекс свободи виборів в Україні є надзвичайно високим за весь період незалежності. Найнижчим цей рівень був на перших виборах в новітній історії України. На нашу думку це зумовлювалось інерцією мислення виборців, а також і кандидатів, внаслідок ще недавнього радянського минулого. Відносно малим цей рівень був і на виборах у 2014 році. Тоді на Україну насувалась неминуча війна з Російською федерацією. Бажання очолити державу у цей трагічний для України час було порівняно небагато. Вибори у два тури були б занадто небезпечними для країни. Відповідальність виборців за долю держави визначила саме такий результат. Один з кандидатів впевнено переміг уже у першому турі.

Найвищим рівень свободи виборів був на останніх виборах у 2019 році. Тут вибори відбувались у два тури. Кількість кандидатів була безпрецедентно високою. Переможець другого туру у першому набрав лише трохи більше тридцяти відсотків голосів виборців.

Якщо ж прослідкувати всі президентські вибори у незалежній Україні, то спостерігається чітка тенденція до зростання індексу свободи виборів (пряма лінія, що визначає довготривалий у часі тренд).

На наступній діаграмі наведені аналогічні результати парламентських виборів в Україні за роки незалежності.

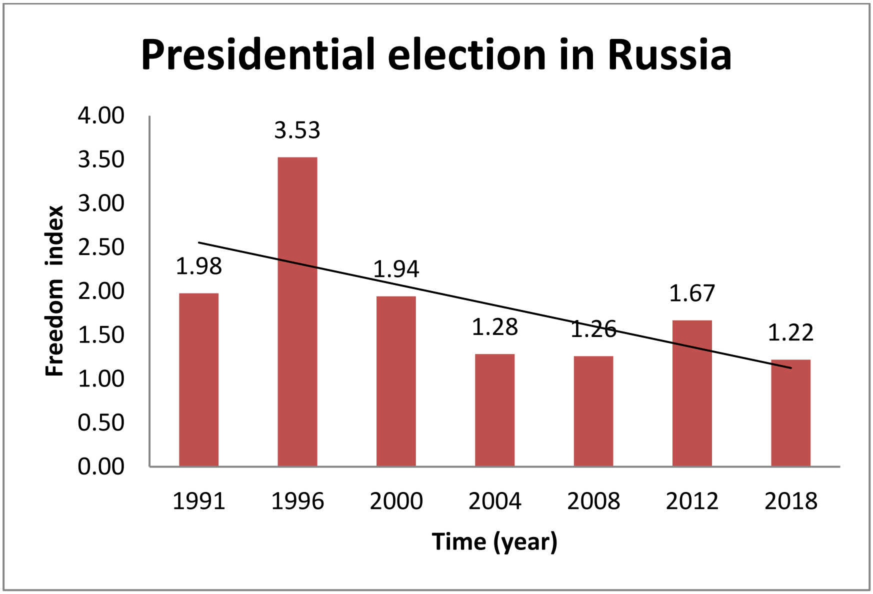


Мал. 4.

Індекс свободи на парламентських виборах також доволі високий і добре корелює з рівнем свободи президентських виборів. Він також має тенденцію до зростання. Найнижчою свобода виборів була на перших парламентських виборах у 1994 році, що узгоджується з найнижчим показником президентських виборів. Проте найвищою вона була у найважчий для України 2014 рік. Реальні шанси на перемогу отримала більша кількість політичних партій, а голоси рівномірніше, ніж на інших парламентських виборах, розподілились між ними. Обравши у травні місяці президента український виборець на патріотичній хвилі завів у парламент величезну кількість національно свідомих депутатів. Зайшли також і партії, що за інших умов не мали такого шансу. У парламенті не з'явилась велика кількість депутатів від окупованих Російською федерацією областей України, традиційно ортогональна до українських цінностей. Ще ніколи парламент не працював так ефективно, як у період з 2014 по 2019 роки.

Президентські і парламентські вибори в Російській федерації

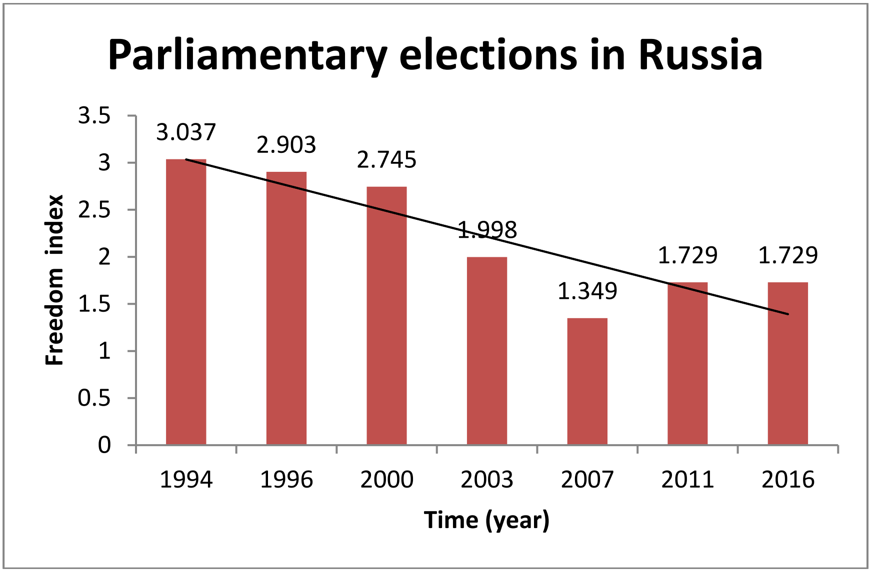
Як ми вже казали, українські вибори найдоцільніше порівнювати з виборами у країнах, що виникли на руїнах Радянського Союзу. Важливим фактором тут є спільність стартових політичних і економічних умов. Також важливим вирівнюючим фактором є менталітет радянської людини, присутній у всіх, навіть найвіддаленіших куточках колишнього Радянського Союзу на момент його розвалу. Якби результатів виборів у різних країнах виявили суттєві відмінності, то ці відмінності, у першу чергу, були б зумовлені відмінностями етнічного походження. Результати президентських виборів у Російській федерації за роки її незалежного існування мають наступний вигляд.



Мал. 5.

З діаграми видно, що рівень свободи президентських виборів у Російській федерації відразу стартував з доволі високого рівня у 1991 році. У 1996 році він досяг максимуму, а останні майже двадцять років він демонструє очевидну тенденцію до зменшення. З математичної точки зору безпосередніми причинами такої поведінки рівня свободи виборів є як зменшення кількості кандидатів, так і збільшення нерівномірності розподілу голосів між ними. Тобто велика кількість кандидатів отримує символічно малу кількість голосів, зате один з кандидатів отримує безпрецедентно високу їх кількість.

Результати парламентських виборів подаються на наступній діаграмі.

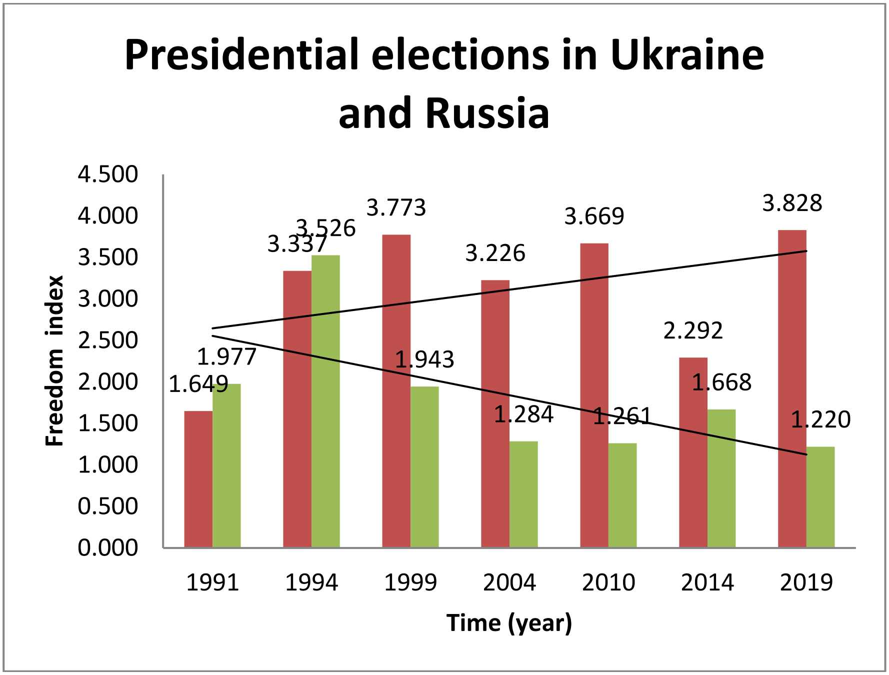


Мал. 6.

Ця діаграма також демонструє високий рівень свободи парламентських виборів у Російській федерації у 1994, 1996, 2000 роках. Проте демонструє також і потужну тенденцію до зменшення цього рівня за всі проаналізовані роки. Спостерігається також сильна кореляція між результатами президентських і парламентських виборів. Така кореляція, як і у разі України, свідчить, на нашу думку, про об'єктивний характер процесу виборів для кожного народу, як би вибори не називались. Тобто характер виборів є відбитком менталітету того, чи іншого народу, його засадничих етнічних цінностей.

Порівняння виборів в Україні і Росії

Тільки порівняння результатів виборів у різних країнах дає можливість визначитись з нашим власним місцем у сучасному світовому політичному процесі. На наступній діаграмі таке порівняння наведене для президентських виборів в Україні і Російській федерації.

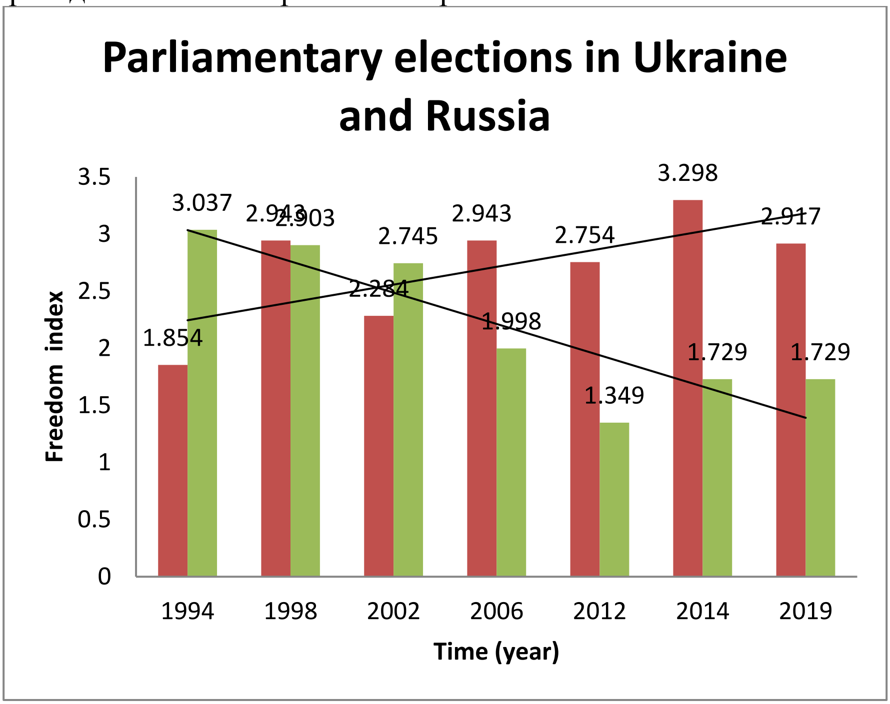


Мал. 7.

З діаграми видно, що рівень свободи виборів у 1991 році стартував в обох країнах практично з одного рівня. Це цілком можна пояснити інерцією мислення українців і мешканців Російської федерації. Покоління виборців в обох країнах переважно сформувались в умовах однієї політичної реальності - Радянському Союзі. Але з часом розбіжності почали наростати. Це відбувалось в міру того, як до виборчих урн в Україні почали приходити покоління виборців, що сформувались політично, або навіть і народились в незалежній українській державі. Відповідно почала суттєво зменшуватись з природних причин кількість виборців, світогляд яких сформувався у Радянському Союзі. Така ж еволюція почалась і серед виборців Російської федерації але у прямо протилежному напрямку. Складається враження, що радянське виховання було певним компромісом для різноманітних етнічних груп, що населяли Радянський Союз.

Розпад Радянського Союзу був одночасно і зникненням цього компромісу. Далі почалась еволюція кожної етнічної групи до свого, характерного саме їй ментально світосприйняття. У висліді ми отримали на останніх президентських виборах в Україні найвищий за всю її історію рівень свободи вибору, а на президентських виборах в Російській федерації, відповідно – найнижчий. При цьому ці результати відрізняються у рази.

На наступній діаграмі ми навели порівняльний аналіз президентських виборів в обох країнах.



Мал. 8.

Свобода парламентських виборів в Україні і Російській федерації також суттєво відрізняться на користь більшої свободи виборів в Україні порівняно з Російською федерацією. Для України спостерігається чітка тенденція до зростання свободи виборів, для Російської федерації – чітка тенденція до її зменшення. Проте є і певні відмінності. Свобода парламентських виборів в Російській федерації стартувала з вищих порівняно з Україною значень. Тенденція до суттєвого зменшення свободи парламентських виборів у Російській федерації стала виразною за останні двадцять років, коли помітно зросла концентрація влади в руках президента.

Висновки

1.Як і у разі президентських виборів, результати всіх останніх парламентських виборів в Україні і Російській федерації суттєво відрізняються між собою на користь більшої свободи виборів в Україні.

2.З наведеного порівняльного аналізу президентських і парламентських виборів в Україні і Російській федерації випливає, на нашу думку, твердження, що суттєва відмінність результатів виборів в обох країнах протягом майже тридцяти років свідчить про суттєву ментальну відмінність українців і мешканців Російської федерації.

3.Індекс свободи виборів в Україні цілком відповідає значенню індексу свободи виборів у провідних країнах Євросоюзу, а інколи навіть і перевищує його. Це, на нашу думку, свідчить про ментальну близькість українців з народами Євросоюзу.

4.Індекс свободи виборів у Російській федерації є суттєво нижчим за значення цього індексу для провідних країн Євросоюзу і України. Це, на нашу думку, свідчить про ментальну відмінність народів Євросоюзу вкупі з Україною, від народів Російської федерації.

5.Всі характерні особливості поведінки індексу свободи виборів, як видно на прикладі двох країн із спільною історією, мають сталий у часі характер і свідчать про наявність глибоких внутрішніх причин такої відмінності.

6.Сама можливість помітити найдрібніші деталі як президентських, так і парламентських виборах в обох країнах на рівні чисел свідчить, на нашу думку, про те, що запропоновані формули для аналізу виборів, є ефективним інструментом кількісного дослідження даного політичного процесу. Ми впевнені, що подібний підхід можна застосувати для кількісного аналізу і інших аспектів політичного життя різних країн.

7.Основні результати роботи надруковані у наступних виданнях: Швець В. Т. Інформаційна ентропія і свобода вибору / Під редакцією С В. Котлика. Інформаційні технології та автоматизація. – Одеса: Астропрінт, 2020, 248 с. Швець В. Т. Ентропія і вибори / Світогляд. – 2019. - № 6. – С. 56 – 61. Shvets V. T. Entropy and freedom of choice / Bintel. – Geopolitical Analitics Journal. – 2019. - № 4. - 94 – 99 p.

4.Флуктуації у фізичних системах

Розглянемо роль флуктуацій в залежності від розмірів системи. У статистичній фізиці величина $D\xi$ носить назву квадратної флуктуації. Якщо $D\xi$ мала, то можливі значення ξ близькі до її середнього значення. Малість $D\xi$ щодо $M\xi$ характеризує величина $\delta\xi = \sqrt{D\xi} / M\xi$, що називається відносною флуктуацією. Зазначеній малості відповідає умова, що $\delta\xi << 1$. .

Важливу роль у статистичній фізиці відіграє наступна теорема. Якщо система складається з N незалежних частин, то відносна флуктуація довільної адитивної функції, що характеризує цю систему, обернено пропорційна кореню з кількості незалежних частин N тобто $\delta\xi \sim 1/\sqrt{N}$ δ

Адитивність випадкової величини ξ означає, що її значення для цілої системи дорівнює сумі значень ξ_k $k = 1,..,N$, де ξ_k у такій же мірі характеризує к-ту незалежну систему як ξ цілу систему, отже

$$\xi = \sum_{k=1}^{N} \xi_k \ .$$

Очевидно, подібна рівність справедлива і для математичних очікувань, тобто

$$M\xi = \sum_{k=1}^{N} M\xi_k \ ,$$

оскільки відповідно до властивості математичного очікування

$$M\sum_{k=1}^{N} \xi_k = \sum_{k=1}^{N} M\xi_k \ ,$$

Обчислимо тепер квадратичну флуктуацію (дисперсію) величину ξ, тобто величину $D\xi$. Для суми незалежних випадкових величин ξ_k має місце наступна властивість дисперсії

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

Тому

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

Якщо випадкові величини ще й однаково розподілені з однаковими математичними очікуваннями і дисперсіями, тобто

$$M\xi_1 = M\xi_2 = ... = M\xi_N = \mu \ ,$$

$$D\xi_1 = D\xi_2 = ... = D\xi_N = \sigma^2 \ ,$$

то

$$M\xi = \mu N \ ,$$

$$D\xi = \sigma^2 N \ .$$

Оскільки

$$\delta\xi = \frac{\sqrt{D\xi}}{M\xi} = \frac{\sigma\sqrt{N}}{\mu N} = \frac{\sigma}{\mu}\frac{1}{\sqrt{N}}$$

Якщо про однаковість розподілу випадкових величин твердження є неточним, то остання рівність стає наближеною.

У фізиці випадковою величиною ξ може бути енергія, імпульс, момент імпульсу, тощо. У соціології – рівень сепаратизму, патріотизму, тощо. Отриманий результат свідчить про наступне. Із зростанням кількості незалежних частин системи N абсолютна флуктуація величини ξ зростає

$$\sqrt{D\xi} \sim \sqrt{N} \,,$$

а відносна надає

$$\delta\xi \sim 1/\sqrt{N} \,.$$

Якщо мова йде про молекули газу і їх кінетичну енергію, то із зростанням їх кількості зростає й ймовірність таких значень кінетичної енергії, що суттєво перевищують її середнє значення для всіх молекул. При цьому вплив таких флуктуацій на енергію всієї системи із зростанням N спадає. Якщо мова йде про сепаратизм, то із зростанням кількості адміністративних одиниць як $\sqrt{N}$ зростає і ймовірність виходу з під контролю країни окремих адміністративних одиниць, але $1/\sqrt{N}$ спадає реальна загроза від такого бунту. За малої кількості адміністративних одиниць ймовірність виходу з під контролю мала, але наслідки такої події руйнівні. Як і завжди – проблема у пошуку золотої середини.

Взявши в якості системи громадян довільної країни, в якості випадкової величини – майно всіх громадян, а в якості випадкових величин ξ_k – майно окремих громадян отримаємо, що ймовірність багатої людини серед громадян країни зростає пропорційно кореню квадратному з кількості громадян, а вплив такої багатої людини на перебіг подій у країні спадає обережно пропорційно тій самій величині.

5.Виграш в азартній грі

Про закон розподілу неприємностей. Продемонструємо знаходження густини розподілу для ідеального газу людей. У цьому разі людей будемо вважати ідентичними, а їх взаємодію між собою зведемо до обміну грошима, що можливо або внаслідок азартної гри, або внаслідок насильства, тощо. Нехай – кількість грошей у довільної людини ξ, а $P_\xi(x)$ – густина розподілу, що описує цю випадкову величину ξ. При цьому

$$dP = P_\xi(x)dx$$

- ймовірність того, що довільна людина має гроші в інтервалі $[x, x+dx]$. Отже перша умова, якій задовольнятиме густина ймовірності – це умова нормування.

$$\int_0^\infty P_\xi(x)dx = 1.$$

Якщо вважати, що кількість грошей окремої людини величина невід'ємна, тобто люди не дають гроші один одному в борг, то нижня границя інтегралу нормування дорівнюватиме нулю, в іншому випадку $-\infty$. При обміні грошима двох людей має виконуватись закон збереження грошей, отже другою умовою буде наступна.

$$x_1 + x_2 = x_3 + x_4.$$

В якості третьої умови візьмемо умову незалежності густини розподілу від часу, тобто умову її стаціонарності. Цю умову можна записати так

$$P_{\xi 1}(x_1)P_{\xi 2}(x_2) = P_{\xi 1}(x_3)P_{\xi 2}(x_4)$$

з наступним поясненням. Двоє людей до зустрічі мали $x_1\, i\, x_2$ грошей. Після зустрічі $x_3\, i\, x_4$ відповідно. Густина функції розподілу двох випадкових величин $\xi_1\, i\, \xi_2$ внаслідок зустрічі не повинна змінитись інакше вона залежала б від часу. Тобто

$$P_{\xi 1 \xi 2}(x_1, x_2) = P_{\xi 1 \xi 2}(x_3, x_4).$$

Оскільки випадкові величини ξ_1 і ξ_2 вважаються незалежними, то

$$P_{\xi 1 \xi 2}(x_1, x_2) = P_{\xi 1}(x_1) P_{\xi 2}(x_4)$$

звідки і отримується наведений вище результат. Умова стаціонарності і є основним рівнянням для знаходження шуканої густини розподілу. Його доцільно об'єднати з законом збереження. У результаті

$$P_{\xi 1}(x_1) P_{\xi 2}(x_2) = P_{\xi 1}(\varepsilon_1) P_{\xi 2}(x_1 + x_2 - x_3)$$

звідки і отримується наведений вище результат. Умова стаціонарності і є основним рівнянням для знаходження шуканої густини розподілу. Його доцільно об'єднати з законом збереження. У результаті

$$P_{\xi 1}(x_1) P_{\xi 2}(x_2) = P_{\xi 1}(x_3) P_{\xi 2}(x_1 + x_2 - x_3).$$

Спростити це функційне рівняння можна лише замінивши його диференційним рівнянням. Для цього логарифмуємо його

$$\ln[P_{\xi 1}(x_1)] + \ln[P_{\xi 2}(x_2)] = \ln[P_{\xi 1}(x_3)] + \ln[P_{\xi 2}(x_1 + x_2 - x_3),$$

а далі здиференцієюємо спочатку за зміною x_1, а тоді - x_2. У результаті

$$\frac{1}{P_{\xi 1}(x_1)} \frac{dP_{\xi 1}(x_1)}{dx_1} = \frac{1}{P_{\xi 2}(x_1 + x_2 - x_3)} \frac{dP_{\xi 2}(x_1 + x_2 - x_3)}{dx_1},$$

$$\frac{1}{P_{\xi 2}(x_2)} \frac{dP_{\xi 2}(x_2)}{dx_2} = \frac{1}{P_{\xi 2}(x_1 + x_2 - x_3)} \frac{dP_{\xi 2}(x_1 + x_2 - x_3)}{dx_2}.$$

Оскільки збігаються криві частини останніх двох рівностей, то збігаються і ліві, тобто

$$\frac{1}{P_{\xi 1}(x_1)} \frac{dP_{\xi 1}(x_1)}{dx_1} = \frac{1}{P_{\xi 2}(x_2)} \frac{dP_{\xi 2}(x_2)}{dx_2}.$$

У цьому рівнянні лівою частиною є функція лише змінної x_1, кривої – лише функція змінної x_2. Дві функції, що залишать від різних змінних збігаються лише у разі, якщо вони є сталими величинами, тобто

$$\frac{1}{P_{\xi 1}(x_1)}\frac{dP_{\xi 1}(x_1)}{dx_1} = \frac{1}{P_{\xi 2}(x_2)}\frac{dP_{\xi 2}(x_2)}{dx_2} = -\alpha,$$

д α - довільна стала. Останню подвійну рівність можна також записати так

$$\frac{1}{P_{\xi}(x)}\frac{dP_{\xi}(x)}{dx} = -\alpha,$$

де $x = x_1, x_2, ...$ Отже отримуємо наступне звичайне диференційне рівняння

$$d\ln[P_{\xi}(x)] = -\alpha dx\,.$$

Його розв’язком є

$$P_{\xi}(x) = A\exp(\alpha x)\,,$$

де A – довільна стала, яку можна знайти з умови нормування,

$$A\int_0^{\infty}\exp(-\alpha x)dx = 1.$$

Звідси $A = \alpha$. Таким чином, шукана густина розподілу має вигляд

$$P_{\xi}(x) = \alpha\exp(-\alpha x)$$

і залежить від єдиного параметра α α. Для його знаходження потрібно врахувати всю суму грошей розподілених між людьми. Для цього обчислимо математичне очікуванням випадкової величини ξ

$$M\xi = \int_0^{\infty} xP_{\xi}(x)dx = \frac{1}{\alpha} = \mu.$$

З іншого боку, відповідно до закону великих чисел.

$$\lim_{N \to \infty} \overline{\xi} = \mu .$$

Тут вибіркове середнє

$$\overline{\xi} = \frac{1}{N} \sum_{k=1}^{N} \xi_k = \frac{E}{N}$$

де Е – загальна сума грошей розподілених між людьми. Отже, якщо кількість людей достатньо велика, то

$$\overline{\xi} = \mu ,$$

або

$$\mu \approx E / N$$

і

$$P_{\xi}(x) \approx \frac{N}{E} \exp(-Nx / E)$$

Отриманий результат має просту інтерпретацію. Якщо N особам видати однакову суму грошей E / N і надалі не втручатись в їх стосунки, то через недовгий час виникне майнова нерівність, що описується експоненційним законом. При цьому найбільша кількість людей повністю втратить все майно ($x = 0$), а незначна частина матиме його дуже багато $x >> E / N$. Цей результат є універсальним і не залежить від якостей людей. Час його досягнення визначається лише інтенсивністю міжлюдських стосунків.

Список літератури:

1.Турчин В. М. Математична статистика. Київ: Видавничий центр "Академія", 1999, 240 с

2.Барковський В. В., Барковська Н. В., Лопатін О. К. Теорія ймовірностей та математична статистика. Київ: ЦУП, 2002, 448 с.

3.Валєєв К. Г., Джалладова І. А. Збірник задач з теорії ймовірностей та математичної статистики. Київ: КНЕУ, 2008, 352с.

4.Мармоза А. Т. Практикум з математичної статистики. Київ: Кондор, 2004, 258 с.

5.Наконечний С.І., Терещенко Т. О., Романюк Т. П. Економетрія. Київ: 2004, КНЕУ, 524 с.

6.Жураковський Ю. П., Полторак В. П. Теорія інформації та кодування. Київ: Вища школа, 2001, 256 с.

7.Тарасов Л. В. Мир построенный на вероятности. Москва: Просвещение, 1984, 92 с.

8.Худсон Д. Статистика для физиков. Москва: Мир, 1970, 296 с.

9.Королюк В. С., Портенко Е. И., Скороход А. В., Турбин А. Ф. Справочник по теории вероятностей и математической статистике. Москва: Наука, 1985, 640 с.

10.Коваленко И. Н., Филиппова А.А. Теория вероятностей и математическая статистика. Москва: Высшая школа, 1982, 256 с.

11.Чистяков В. П. Курс теории вероятностей. Москва: Наука, 1987, 240 с.

12.Боровков А. А. Теория вероятностей. Москва: Наука, 1976, 352с.

13.Феллер В. Введение в теорию вероятностей и приложения. Том 1. Москва: Мир, 1984, 528 с.

14.Феллер В. Введение в теорию вероятностей и приложения. Том 2. Москва: Мир, 1984, 748 с.

14.Рудавський Ю. А., Понеділок Г. В. Функціональні інтеграли, та їх застосування. Львів: Львівська політехніка, 2002, 314 с.
15.Кособуцький П. С., Лобур М. В. Статистичне моделювання. Львів: Львівська політехніка, 2013, 328 с.