

В.Т. Швець

**Теорія ймовірностей і
математична статистика**

Одеса 2018

ББК 22.161я73
УДК 517(0.75.8)

Швець Валерій Тимофійович
Теорія ймовірностей і математична статистика
Одеса. Видавництво ВМВ, 2018 - 218 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 ймовірностей та математичної статистики 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 техн. наук, проф. В.Х.Кирилов

© В.Т.Швець 2018

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 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, математичних очікувань тощо. Тому роль таких наук в математичній підготовці студентів технічних вузів як: теорія ймовірностей, математична статистика, випадкові процеси має невинно зростати. Аудиторія тих, кому потрібні ці знання за межами навчальних закладів, також має невинно зростати і вона 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.

Передмова	3
Глава 1	6
Елементи теорії ймовірностей	6
1. 1. Частотне означення ймовірності	6
1. 2. Аксиоматичне означення ймовірності	9
1. 3. Елементи комбінаторики	12
1. 4. Класичне означення ймовірності	18
1. 5. Геометричне означення ймовірності	24
1. 6. Формула повної ймовірності	34
1. 7. Послідовність незалежних випробувань	39
1.8. Інформація та ентропія	45
1. 8. Задачі для самостійного розв'язання	52
Глава 2.	58
Випадкова величина	58
2.1. Випадкова величина	58
2.2. Закон розподілу кількох випадкових величин	62
2.3. Закон розподілу для функції неперервної випадкової величини	67
2.4. Математичне очікування	69
2.5. Дисперсія	76
2.6. Коваріація. Коефіцієнт кореляції	89
2.7. Моменти випадкової величини	91
3.4. Твірна функція моментів	94
2.8. Класичний ідеальний газ	97
2.9. Закон великих чисел	106
2.10. Граничні теореми в схемі Бернуллі	109
2. 11. Задачі для самостійного Розв'язання	119
Глава 3	125
Елементи математичної статистики	125
3.1. Вибірка	125
3.2. Вибіркові моменти	132

3.5. Вибіркові моменти для великих вибірок	137
3.6. Емпірична функція розподілу	143
3.7. Оцінювання параметрів розподілу	147
3.8. Точкова оцінка	149
3.8.1. Метод найбільшої правдоподібності	157
3.8.2. Метод моментів	160
3.9. Інтервальні оцінки	163
3.9.1. Довірчий інтервал для математичного очікування нормального розподілу при відомій дисперсії	164
3.9.2. Довірчий інтервал для математичного очікування нормального розподілу при невідомій дисперсії	170
3.9.3. Довірчий інтервал для дисперсії нормального розподілу при відомому математичному очікуванні	176
3.10. Статистична перевірка гіпотез.	
Критерій Пірсона	180
3.11. Статистична перевірка гіпотези про математичне очікування нормально розподіленої генеральної сукупності	188
3.12. Статистична перевірка гіпотези про дисперсію нормально розподіленої генеральної сукупності	194
3.13. Регресійний аналіз	198
3.14. Метод найменших квадратів	203
3.15. Дисперсійний аналіз	205
3. 16. Задачі для самостійного розв'язання	208
Список літератури:	227

Глава 1. Елементи теорії ймовірностей

1.1. Частотне означення ймовірності

Найважливішим об'єктом теорії ймовірностей є сама 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. Якщо ми підкидаємо монету десять разів, то провадимо десять випробувань. В кожному з них можуть спостерігатись лише дві події: випадення гербу 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. Отже, в теорії ймовірностей кожній події ставиться у відповідність її ймовірність. Ймовірність є абстрактним поняттям, суть якого можна зрозуміти, пов'язавши його з простішою числовою характеристикою події, що називається частотою події. Остання просто означає долю від

загальної кількості випробувань в якій спостерігається подія, що нас цікавить, і 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, оскільки залежить від N . Також вона може бути іншою і 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

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

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 $n_N(A+B) = n_N(A) + n_N(B)$, відповідно

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

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

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

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 більшим 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 оснований на властивості межового переходу

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

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

1.2. Аксиоматичне означення ймовірності

На завершальному етапі свого розвитку будь-яка математична 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, або витвори чистого розуму. 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, впливають 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 $0 \leq P(A) \leq 1$. При цьому ймовірність неможливої події - пустої множини $P(\emptyset) = 0$, ймовірність достовірної події - події, підмножинами якої є всі можливі випадкові події даного випадкового експерименту, $P(\Omega) = 1$. Така максимально можлива множина називається ще простором елементарних подій.

2. Її областю визначення є всі можливі події, що характеризують даний випадковий експеримент разом з

неможливою і достовірною подіями. Останнє твердження можна і підсилити. Областю визначення ймовірності є алгебра подій U , тобто множина всіх подій, замкнена відносно всіх операцій над подіями: додавання, віднімання, множення і заперечення. Якщо з якихось причин неможливо розглянути максимальну кількість подій, то фактично розглянута сукупність подій має утворювати алгебру. Розглядають також σ -алгебру подій. Це, коли множина подій замкнена відносно зліченої кількості операцій над подіями.

3. Властивість адитивності ймовірності скінченної суми несумісних подій

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

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

4. Властивість мультиплікативності ймовірності скінченного добутку незалежних подій

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

Цю властивість також можна підсилити, поширивши її на нескінченний добуток незалежних подій.

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

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

де A_n - спадна послідовність подій така, що $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$

$$\text{і } \bigcap_{n=1}^{\infty} A_n = \emptyset .$$

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 визначень на ній ймовірності, (Ω, U, P) прийнято називати ймовірнісним простором. Якщо простір елементарних подій складається із

скінченної або зліченої кількості елементарних подій, то такий ймовірнісний простір називається дискретним. Якщо елементарні події утворюють континуум, то неперервним. Аксиома $P(\Omega) = 1$ фактично є умовою нормування ймовірності. Для дискретного ймовірнісного простору $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ остання має вигляд

$$\begin{aligned} P(\Omega) &= P(\{\omega_1, \omega_2, \dots, \omega_n, \dots\}) = \\ &= P(\omega_1 + \omega_2 + \dots + \omega_n + \dots) = \\ &= P(\omega_1) + P(\omega_2) + \dots + P(\omega_n) + \dots = 1. \end{aligned}$$

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

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

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

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

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

$$\begin{aligned} P(A + B) &= P(A) + P(B\bar{A}), \\ P(B) &= P(AB) + P(B\bar{A}). \end{aligned}$$

Виключаючи звідси ймовірність $P(B\bar{A})$ 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

$$P(AB) = P(A)P(B|A).$$

Ця формула не доводиться. Вона є узагальненням четвертої аксіоми ймовірності у разі залежних подій і, фактично, є означенням функції $P(B|A)$ - умовної ймовірності події B , тобто ймовірності події B при умові, що подія A відбулась.

У разі трьох довільних подій ця формула матиме вигляд

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

Тут $P(C|AB)$ 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

$$A = (x_1, x_2, \dots, x_n).$$

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

$$B = \{y_1, y_2, \dots, y_n\}.$$

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

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

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

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

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

$$A_n^m = n(n-1)\dots(n-m+1) = \frac{n!}{(n-m)!}, \quad 0 \leq m \leq n,$$

де $0!=1$, $n!=1\cdot2\cdot\ldots\cdot n$, $n\in N$. Зокрема, $A_n^0=1$, $A_0^0=1$. Якщо $M=\{1,2,3\}$ – триелементна неупорядкована множина, то з її елементів можна утворити $A_3^1=3$ розміщень з трьох по одному елементу: (1) , (2) , (3) ; $A_3^2=3\cdot2=6$ розміщень з трьох по два елементи: $(1,2)$, $(2,1)$, $(1,3)$, $(3,1)$, $(2,3)$, $(3,2)$; $A_3^3=3\cdot2\cdot1=6$ 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 неупорядковану m елементну множину, утворену з неупорядкованої n – елементної множини. Кількість сполучень визначається формулою

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

Зокрема, $C_n^0=1$, $C_n^n=1$. Якщо $M=\{1,2,3\}$ – триелементна неупорядкована 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 $(x_1, x_2, \dots, x_k)$, 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) , $(3,3)$, $(3,3,3)$, ..., $(1,2)$, $(1,1,2)$,

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

Дві впорядковані вибірки називаються рівними, якщо вони мають однакові довжини та однакові елементи на однакових місцях. Дві неупорядковані вибірки називаються рівними, якщо вони мають однакові довжини і кожний їх елемент повторюється однаково 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 неупорядкованими, 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

$$A_n^m = n^m.$$

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 $A_3^2 = 3^2 = 9$.

Перестановками з повтореннями з n елементів по m називають упорядковані вибірки довжиною m , утворені з n –

елементної множини n , при умові що $m = m_1 + m_2 + \dots + m_n$, 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

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

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 $P_3 = 3!/3!0! = 1$. 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 $P_3 = 3!/2!1! = 3$.

Сполученнями з повтореннями з n елементів по m називають неупорядковану 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 $C_3^2 = C_4^2 = \frac{4!}{2!2!} = 6$.

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, $78 = C_{n-3}^2$, тобто

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

або

$$n^2 - 7n - 144 = 0.$$

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

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

Нехай всі елементарні події рівноправні, тобто

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

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

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

З тих же міркувань, що і вище

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

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

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

Позначаючи кількість елементів скінченної 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. Ще раз нагадуємо, що останнє можливо лише у разі рівноправності всіх її елементів.

Розглянемо декілька прикладів.

Задача 4. Дитина, граючись десятьма кубиками, на яких написані літери М, М, Т, Т, А, А, А, К, И, Е, 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 А можна розташувати на трьох місцях $P_3 = 3!$ способами; кубики з літерою М можна розташувати на двох місцях $P_2 = 2!$ способами, так само як і кубики з літерою Т – $P_2 = 2!$ способами. Тоді подія А, що відповідає утворенню з літер на кубиках слова математика, складається з $|A| = 3! \cdot 2! \cdot 2! = 24$ подій. Відповідно до класичного означення ймовірності

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

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

Задача 5. 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 $P_{n-1} = (n-1)!$, оскільки одне місце зайняте потрібним ключем, а решта $n-1$ ключів можна на $n-1$ місцях розкласти $|A| = P_{n-1} = (n-1)!$ способами. Таким чином шукана ймовірність дорівнює

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

Задача 6. 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 $A = \{\omega_1, \omega_4\}$, що нас цікавить, складається лише з двох елементарних подій $|A| = 2$.
Відповідно до класичного означення ймовірності

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

Задача 7. У колоді 36 карт. Знайти ймовірності подій:

$A = \{\text{чотири тузи розташовані поряд}\}$,

$B = \{\text{місця розміщення тузів утворюють арифметичну прогресію з кроком 7}\}$.

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 8. Кинуті три монети. Вважаючи всі елементарні події рівноправними, знайти ймовірності подій:

$$A = \{\text{перша монета випала гербом догори}\},$$

$$B = \{\text{випало рівно два герба}\},$$

$$C = \{\text{випало не більше двох гербів}\}.$$

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 9. З множини всіх послідовностей довжини n , що складаються з цифр 0, 1, 2, випадково обирається одна. Знайти ймовірності подій:

$A = \{\text{послідовність починається з нуля}\},$

$B = \{\text{послідовність містить рівно } m+2 \text{ нуля, причому два з них знаходяться на кінцях послідовності}\},$

$C = \{\text{послідовність містить рівно } m \text{ одиниць}\},$

$D = \{\text{у послідовності рівно } m_0 \text{ нулів, } m_1 \text{ одиниць, } m_2 \text{ двійок}\}.$

Розв'язання. Елементарною подією вважатимемо довільну послідовність з цифр. Оскільки кожний елемент цієї послідовності може бути однією з трьох цифр, то загальна кількість елементарних подій дорівнюватиме $|\Omega| = 3^n$. У разі події A перший елемент цієї послідовності зафіксований, а решта довільні. Отже, загальна кількість елементарних подій, з яких складається подія A , дорівнюватиме $|A| = 3^{n-1}$. У висліді

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

Аналогічно

$$P(B) = \frac{C_{n-2}^m 2^{n-m-2}}{3^n},$$

$$P(C) = \frac{2^{n-m} C_n^m}{3^n},$$

$$P(D) = \frac{n!}{3^n m_0! m_1! m_2!}.$$

Наступна задача пов'язана з розглядом простору елементарних подій, що містить злічену кількість елементарних подій.

Задача 10. Нехай симетрична монета підкидається доти, доки не випаде герб. Знайти ймовірність того, ця подія

відбудеться при n -'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 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$. Ймовірність того, що зазначена подія відбудеться при n -ому підкиданні є $P(\omega_n) = (1/2)^n$. При цьому,

очевидно, виконується умова нормування $\sum_{n=1}^{\infty} P(\omega_n) = \sum_{n=1}^{\infty} (1/2)^n = 1$.

Отже, ймовірність того, що це відбудеться на n -'ятому підкиданні є $P(\omega_5) = (1/2)^5 = 1/32$, за парне число підкидань

$\sum_{n=1}^{\infty} P(\omega_{2n}) = \sum_{n=1}^{\infty} (1/2)^{2n} = 1/3$, 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 такою зручною числовою характеристикою є кількість елементів множини. У разі рівноправності елементів множин, ця числова характеристика може слугувати мірою таких множин, оскільки множини можна легко порівнювати між собою. У разі нерівноправних елементів кількість елементів множин мало про що говорить. Значно зручнішим тут 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 в якості зручної числової характеристики речовини може виявитись необхідним при підсумовуванні враховувати нескінчену кількість доданків (дивись останню задачу). Для нескінчених множин, елементи яких утворюють неперервну множину, тобто множину, між будь-якими елементами якої обов'язково містяться інші елементи цієї ж множини, слід обрати іншу числову характеристику. В цьому разі немає сенсу для порівняння використовувати таку характеристику як кількість елементів множин. Аналогічно випадку дискретних множин, кожному елементу множині (кожній точці) доцільно поставити у відповідність число. Сукупність всіх таких чисел є нічим іншим як функцією неперервного аргументу. Серед нескінченної 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 (густин розподілу ймовірностей), які матимуть сенс певних 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 $m(A)$, хоча також вживатимемо і позначення $|A|$, і надаючи йому універсального змісту, матимемо наступний аналог класичного значення ймовірності

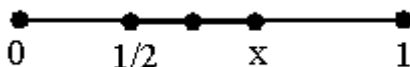
$$P(A) = \frac{m(A)}{m(\Omega)},$$

або

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

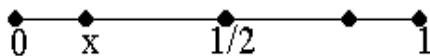
Задача 11. Випадкова точка рівномірно розподілена на відрізку $[0,1]$ і ділить цей відрізок на дві частини. Нехай η_1 – довжина більшої частини і η_2 – довжина меншої частини. Знайти $P\{\eta_1 \leq x\}$, $P\{\eta_2 \leq x\}$ при довільному x .

Розв’язання. Оскільки η_1 – довжина більшої частини відрізку $[0,1]$, то умова $\eta_1 \leq x$ може реалізуватись лише у разі, якщо $x \in [1/2, 1]$. Вважаючи елементарною подією довільну точку проміжку $[0,1]$, маємо $|\Omega|=1$. Елементарні події, з яких складається подія $A(\eta_1 \leq x)$ належать наступній множині $[1-x, x]$. Довжина цього



Мал. 1.

відрізку $|A| = x - (1-x) = 2x - 1$. Отже, $P(\eta_1 \leq x) = 2x - 1$. Оскільки η_2 – довжина меншої частини відрізку $[0,1]$, то умова $\eta_2 \leq x$ може реалізуватись лише у разі, якщо $x \in [0, 1/2]$. У цьому разі подія $B(\eta_2 \leq x)$ складається з елементарних подій, що



Мал. 2.

належать проміжкам $[0, x] \cup [-x, 1]$. Їх загальна площа $|B| = x + [1 - (1 - x)] = 2x$. Отже, $P(\eta_2 \leq x) = 2x$.

Задача 12. Випадкова точка A має рівномірний розподіл у квадраті зі стороною 1. Знайти ймовірності наступних подій:
 $A = \{\text{відстань від точки } A \text{ до фіксованої сторони квадрата не перевищує } x\}$,

$B = \{\text{відстань від точки } A \text{ до найближчої сторони квадрата не перевищує } x\}$,

$C = \{\text{відстань від точки } A \text{ до центра квадрата не перевищує } x\}$,

$D = \{\text{відстань від точки } A \text{ до фіксованої вершини квадрата не перевищує } x\}$.

Розв'язання.

а)

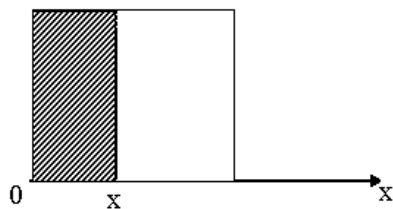
Мал. 3.

$$|\Omega|=1,$$

$$|A|=1 \cdot x = x$$

$$|A|=1,$$

Отже,



$$, x \leq 1,$$

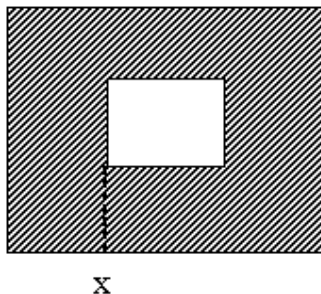
$$1 < x < \infty.$$

$$P(A) = \min(x, 1),$$

$$P(A) = x, \quad 0 \leq x \leq 1,$$

$$P(A) = 1, \quad x \geq 1.$$

6)



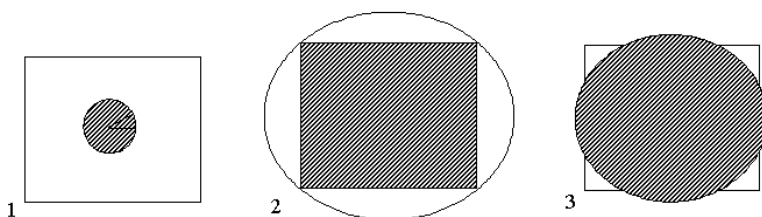
Мал. 4.

$$|\Omega|=1, \quad |B|=1-(1-2x)^2=4x(1-x)$$

$$P(B)=4x(1-x), \quad 0 \leq x \leq 1/2$$

$$P(B)=1, x \geq 1/2$$

c)



Мал. 5.

$$|\Omega|=1, |C|=\pi x^2, P(C)=\pi x^2, 0 \leq x \leq 1/2$$

$$|\Omega|=1, |C|=1, P(C)=1, x \geq \sqrt{2}$$

$$|\Omega|=1, |C|=x^2 \left[\pi - 4a \cos(1/2x) \right] + \sqrt{4x^2 - 1}, 1/2 \leq x \leq 1/\sqrt{2}$$

$$P(C)=x^2 \left[\pi - 4a \cos(1/2x) \right] + \sqrt{4x^2 - 1}$$

d)

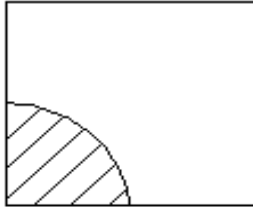
Мал. 6.

$$P(D) = \frac{1}{4} \pi x^2,$$

$$P(D) = 1,$$

$$0 \leq x \leq 1$$

$$x \geq \sqrt{2}$$

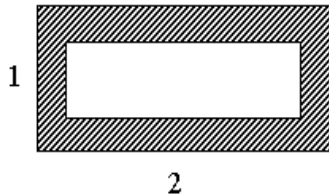


$$P(D) = x^2 \left(\frac{\pi}{4} - \arccos \frac{1}{x} \right) + \sqrt{x^2 - 1}, \quad 1 \leq x \leq \sqrt{2}.$$

Задача 13. Випадкова точка A має рівномірний розподіл у прямокутнику з сторонами 1 і 2. Знайти ймовірності наступних подій: $A = \{\text{відстань від } A \text{ до найближчої сторони прямокутника не перевищує } x\}$, $B = \{\text{відстань від } A \text{ до довільної сторони прямокутника не перевищує } x\}$, $C = \{\text{відстань від } A \text{ до діагоналей прямокутника не перевищує } x\}$.

Розв'язання.

а)



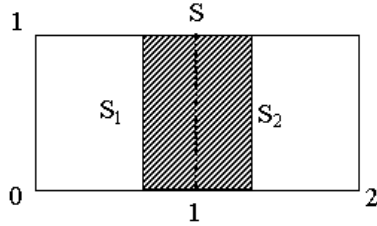
Мал. 7.

$$|\Omega| = 2, \quad |A| = 2 - (2 - 2x)(1 - 2x) = 2x(3 - 2x)$$

$$P(A) = x(3 - 2x), \quad 0 \leq x \leq 1/2$$

$$P(A)=1, \quad x \geq 1/2$$

б)



Мал. 8.

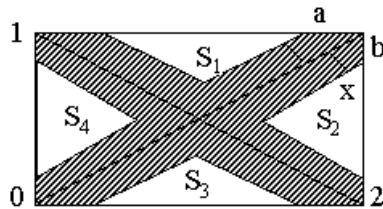
$$|\Omega|=2, \quad |B|=S-2S_1, \quad S_1=2-x, \quad S=2,$$

$$P(B)=\frac{2-(2-2x)}{2}=x-1, \quad 1 \leq x \leq 2,$$

$$P(B)=1, \quad x \geq 2,$$

$$P(B)=0, \quad 0 \leq x \leq 1.$$

с)



Мал. 9.

$$|\Omega|=2, \quad a=x\sqrt{5}, \quad b=a/2=x\sqrt{5}/2,$$

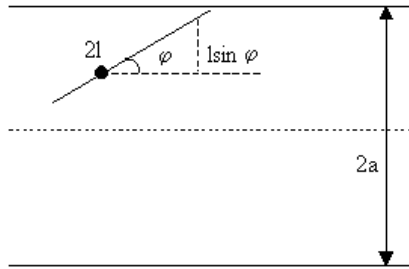
$$S_1+S_2+S_3+S_4=(2-2x\sqrt{5}) \times (1-x\sqrt{5})=2(1-x\sqrt{5})^2$$

$$|C|=2-2(1-x\sqrt{5})^2, \quad P(C)=1-(1-x\sqrt{5})^2, \quad 0 \leq x \leq 1/\sqrt{5}$$

$$P(C)=1, \quad x \geq 1/\sqrt{5}.$$

Задача 14. (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.

Розв'язання. Нехай n – відстань від центру голки до



Мал. 10.

найближчої прямої, а φ – кут між голкою і цією прямою. Пара чисел (n, 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

$$\Omega = \{(n, \varphi) : 0 \leq \varphi \leq \pi, 0 \leq n \leq a\}.$$

Перетин голки з прямою відбувається лише у тому разі, коли $n \leq l \sin \varphi$. Таким чином, подія A , що нас цікавить, має вигляд

$$A = \{(n, \varphi) : 0 \leq n \leq l \sin \varphi, 0 \leq \varphi \leq \pi\}.$$

Так як

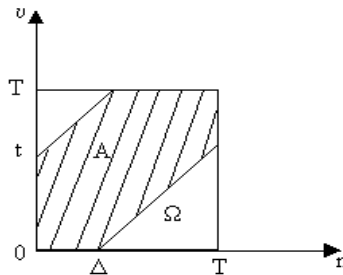
$$|A| = \int_0^{\pi} l \sin \varphi d\varphi = 2l, \quad |\Omega| = \pi \cdot a,$$

то

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

Задача 15. За проміжок часу $[0, T]$ у випадковий момент n з'являється сигнал довжиною Δ . Приймач вмикається у випадковий момент $v \in [0, T]$ на час t . Вважаючи, що точка (n, v) рівномірно розподілена в квадраті $[0, T] \times [0, T]$, знайти ймовірність виявлення сигналу.

Розв'язання. У якості елементарної події слід взяти довільну точку (n, v) , а в якості простору елементарних подій – квадрат $\Omega = \{(n, v) : 0 \leq n \leq T, 0 \leq v \leq T\}$. Якщо сигнал з'являється в момент часу $n = 0$, то він може бути виявленим, якщо приймач вмикається у будь – який момент проміжку $[0, t]$. Якщо сигнал



з'являється у момент часу n ($0 \leq n \leq \Delta$), то він може бути виявленим, якщо приймач вмикається у будь – який момент проміжку $[0, t+n]$. Якщо сигнал з'являється у момент часу ($\Delta < n \leq T$), то він може бути виявленим, якщо приймач вмикається у будь – який момент проміжку $[n, t+n]$. Таким чином, подія A , що нас цікавить, має вигляд

Мал. 11.

Тоді

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

Відповідно,

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

1.6. Формула повної ймовірності

Класичне і геометричне означення ймовірності вичерпують собою алгоритми безпосереднього обчислення ймовірностей. В усіх інших випадках або доцільно, або безваріантно необхідно використання складніших алгоритмів обчислення ймовірностей. Типовим для теорії ймовірностей є структуризація випадкового експерименту, зокрема, розбиття його на незалежні або квазінезалежні 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 $B_1, B_2, \dots, B_n$ попарно несумісні і такі, що $B_1 + B_2 + \dots + B_n = \Omega$. При цьому вважатимемо, що $P(B_k) > 0$, $k = 1, 2, \dots, n$ і A має ненульовий перетин з кожною з множин $B_1, B_2, \dots, B_n$. Тоді матиме місце формула

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

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

$$A = A * B_1 + A * B_2 + \dots + A * B_n.$$

Тоді, на основі згаданої аксіоми,

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

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

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

$$P(A)P(B_k | A) = P(B_k)P(A | B_k),$$

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

$$P(A * B_k) = P(B_k)P(A | B_k),$$

$$P(A * B_k) = P(A)P(B_k | A).$$

З зазначеної тотожності отримуємо формулу Баєса

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

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

Розглянемо декілька задач на використання формули повної ймовірності.

Задача 16. 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 добрий білет, подія B – другий студент взяв добрий білет, подія C – обидва студенти взяли добрі білети. 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 добрий білет

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

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

$$P(B/\bar{A}) = \frac{|B/\bar{A}|}{|\Omega|} = \frac{5}{24} = \frac{5}{24},$$

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

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

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

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

Ймовірність того, що обидва студенти взяли добрий білет 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 17. Задача. 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.

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

б) Знайти ймовірність того, що переливання можна здійснити, якщо є два донори; три донори.

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

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

Перейдемо тепер до пункту б). Нехай подія A_n
 ($n=1,2,3,4$) полягає у тому, що у першого донора n – та група
 крові; подія B_n – відповідає тому, що у другого донора n – та
 група крові; C – у висліді переливання хворий виживе. У цьому разі
 повна система гіпотез має 16 гіпотез $H_{nm} = A_n \cdot B_m$ ($n, m=1,2,3,4$)
 . Подія H_{nm} полягає у тому, що у першого донора n – та група
 крові, а у другого m – та. Оскільки події A_n і B_m незалежні, то

$$P(H_{nm}) = P(A_n \cdot B_m) = P(A_n)P(B_m).$$

Формула повної ймовірності має вигляд

$$\begin{aligned} P(C) &= \sum_{n,m=1}^4 P(H_{nm})P(C/H_{nm}) = \\ &= \sum_{n=1}^4 P(A_n)P(B_m)P(C/A_n \cdot B_m). \end{aligned}$$

Подальші обчислення очевидні. Наприклад

$$P(C/A_1B_2)=1,$$

оскільки наявність донора з першою групою крові гарантує
 виживання хворого з довільною групою крові;

$P(C/A_2B_3)=1-0.337=0.663$, оскільки друга або третя групи крові не влаштовують лише хворого з першою групою крові.

У разі трьох донорів формула повної ймовірності матиме вигляд

$$P(D) = \sum_{n,m,l=1}^4 P(A_n)P(B_m)P(C_l)P(D/A_n \cdot B_m \cdot C_l).$$

Тут подія A_n – у першого донора n – та група крові, B_m – у другого донора m – та група крові, C_l – у третього донора l – та група крові, D – хворий в наслідок переливання виживе.

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 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} . Можна ж 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.

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 $m_1 + m_2 + \dots + m_N = n$, матиме вигляд

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

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

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

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

Остання формула називається формулою Пуассона.

Інша асимптотична формула виникає, якщо $n \rightarrow \infty$ але при цьому p довільне. Тоді має місце локальна формула Муавра-Лапласа

$$P_n(m) \approx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_m^2}{2}\right),$$

де у правій частині наближеної рівності 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 кількості успіхів певному інтервалу ($a \leq m \leq b$), то за тих же умов має місце інтегральна формула Муавра-Лапласа

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

Останній результат можна виразити через функцію Лапласа

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

так

$$P(a \leq x \leq b) = \Phi_0(b) - \Phi_0(a)$$

або через функцію нормального розподілу випадкової величини x у подібний спосіб

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

а саме

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

Очевидно, функція Лапласа є непарною

$$\begin{aligned} \Phi_0(-x) &= \frac{1}{\sqrt{2\pi}} \int_0^{-x} \exp(y^2/2) dy = \\ &= -\frac{1}{\sqrt{2\pi}} \int_0^x \exp(y^2/2) dy = -\Phi_0(x), \end{aligned}$$

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

Великого практичного значення цим формулам надавали в докомп'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

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

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 $P(|m/n - p| < \Delta)$. Структура випадкової величини $|m/n - p|$ не дозволяє застосувати безпосередньо до неї теорему Муавра-Лапласа. Але візьмемо до уваги, що випадкова величина $x_m = \frac{m - np}{\sqrt{np(1-p)}}$, 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

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

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

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

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

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 послуговуватись лише оцінкою. Її легко отримати з тих міркувань, що $p(1-p) \leq 1/4$. Тоді $1/p(1-p) \geq 4$, а

$$\Phi_0 \left(\Delta \sqrt{\frac{n}{p(1-p)}} \right) \geq \Phi_0 \left(2\Delta\sqrt{n} \right). \text{ Шукана оцінка матиме вигляд}$$

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

Задача 18. Вважаючи $p=1/2$, визначити, скільки підкидань монети слід здійснити, щоб із ймовірністю не менше ніж 0.9 відносна частота m/n відрізнялась від $1/2$ на 0.05.

Розв'язання. Скористаємось нерівністю

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

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

Задача 19. Нехай N – досить велика кількість однотипних виробів, з яких M – браковані. Тоді природно вважати, що навімання 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: навімання 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 $\Phi_0(x)$ знайдемо, що $\Phi_0\left(2 \cdot 0.05 \cdot \sqrt{10000}\right)\Phi_0(1)=0.3413$.

Отже,

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

1. 8. Інформація та ентропія

Дуже часто в математиці базою для створення математичної основи нової математичної теорії є добре розвинутий апарат вже якоїсь існуючої теорії. Так, на диво плідним в теорії ймовірностей виявилось використання теорії

множин. Базовий об'єкт теорії ймовірностей найзручніше виявилось визначати як числову функцію множини. При цьому дії над множинами відповідним чином трансформувались у дії над такими числовими функціями. У свою чергу, теорія інформації може цілком запозичити свій математичний апарат у теорії ймовірностей. Просто в теорії ймовірностей ми цікавимося такою її числовою характеристикою як ймовірність, а в теорії інформації – такою її характеристикою як кількість інформації. Кількість інформації розглядатиметься нами як дискретна випадкова функція визначена на просторі елементарних подій, а далі як неперервна.

Отже, розглянемо дві події A і B , в теорії інформації їх прийнято називати повідомленнями, що належать певному ймовірнісному простору. Порівняємо кількості інформації, що містяться в цих подіях. За означенням, кількість інформації $I(A|B)$, що міститься в події A відносно події B , визначається так

$$I(A|B) = \log \left(\frac{P(A|B)}{P(A)} \right),$$

де $P(A)$ - ймовірність події A (ймовірність появи повідомлення A), $P(A|B)$ - умовна ймовірність події A при умові, що наступила подія B (ймовірність появи повідомлення A при умові, що перед тим ми отримали повідомлення B). Введена таким чином функція є симетричною щодо своїх аргументів

$$I(A|B) = I(B|A).$$

Фактично ця функція характеризує взаємовплив двох подій в плані інформації. Наприклад, якщо ці події незалежні, тобто інформація про одну з них не містить жодної інформації про можливість іншої, то $I(A|B) = I(B|A) = 0$

Якщо $B=A$, при цьому $P(A|B)=1$, то ми вже отримуємо можливість вимірювання абсолютної інформації про окрему подію. За означенням, кількість інформації, що міститься в події (повідомленні) A визначається так

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

Очевидно, що кількість інформації не може бути від'ємною величиною. Подібне означення кількості інформації було введено К.Е. Шенноном у 1948 році, а для деяких частинних випадків ще у 1928 році Р. Хартлі. З означення кількості інформації випливає, що чим менша ймовірність події, тим більшою є кількість інформації, що в ньому міститься. Повідомлення про те, що літак x приземлився мало кого зацікавить, оскільки ймовірність такого ходу подій цілком очікувана і відбувається з ймовірністю, дуже близькою до одиниці. Повідомлення про те, що літак x розбився при приземленні зацікавить майже всіх, оскільки ймовірність такого ходу подій близька до нуля, тобто є практично неочікуваною. Очевидно, що

$$\lim_{P(A) \rightarrow 0} I(A) = - \lim_{P(A) \rightarrow 0} \log[P(A)] = \infty,$$

$$\lim_{P(A) \rightarrow 1} I(A) = - \lim_{P(A) \rightarrow 1} \log[P(A)] = 0.$$

Якщо інформацію вимірювати в бітах і вимагати, щоб події (повідомленню) про те, що при одноразовому підкиданні симетричної монети випав герб відповідала кількість інформації в один біт, то логарифм в останньому виразі слід брати по основі 2, що є загальноприйнятим в теорії інформації. Але можна використати і безліч інших варіантів для основи логарифму.

Якщо ви ставите фігуру на довільну клітину C шахової дошки, то ви обираєте один варіант з 64 рівноправних. Відповідно до класичного означення ймовірності ймовірність зробити це навмання $P(C) = 1/64$. Передавши повідомлення про цю подію ви надали іншій особі наступну кількість інформації

$$I(C) = -\log[P(C)] = -\log(1/16) = 6,$$

тобто 6 біт інформації.

З означення кількості інформації і теореми множення ймовірностей випливає, що для незалежних подій A і B має місце умова адитивності інформації, тобто

$$\begin{aligned} I(AB) &= -\log[P(AB)] = -\log[P(A)P(B)] = \\ &= -\log[P(A)] - \log[P(B)] = I(A) + I(B). \end{aligned}$$

В попередній задачі подію C можна собі уявити як вислід одночасного відбування двох інших подій: ви поставили фігуру у

рядок A і стовпець B . У цьому разі ви повідомити іншій людині наступну кількість інформації

$$I(C) = I(AB) = I(A) + I(B),$$

де події A і B , очевидно, незалежні. Якщо вважати всі рядки і стовпці шахівниці рівноправними з точки зору встановлення на них даної фігури і скористатися класичним означенням ймовірності, то $P(A) = 1/8$, $P(B) = 1/8$ і

$$I(C) = 3 + 3 = 6,$$

Тобто ми отримали попередній результат у інший спосіб.

Якщо події A і B довільні, то

$$\begin{aligned} I(AB) &= -\log[P(AB)] = -\log[P(A)P(B|A)] = \\ &= -\log[P(A)] - \log[P(B|A)] = -\log[P(A)] - \log\left[\frac{P(B|A)}{P(B)}P(B)\right] = \\ &= -\log[P(A)] - \log[P(B)] - \log\left(\frac{P(B|A)}{P(B)}\right) = \\ &= I(A) + I(B) - I(A|B) \end{aligned}$$

і спостерігається інтерференція подій (повідомлень), дещо подібна до теореми додавання ймовірностей.

Припустимо, що лише половина запущених торпед потрапляє у ціль. З них лише половина є такими, що топлять корабель. У висліді торпедної атаки корабель затонув. Яку інформацію містить таке повідомлення? Очевидно, що лише одна четверта запущених торпед здатна потопити корабель. Тобто ймовірність такої події $P(C) = 1/4$, а відповідна інформація повідомлення $I(C) = 2$. Тут також подію C можна розглянути як добуток двох подій: події A - корабель вражений торпедою і події B - корабель внаслідок такого враження затонув. Очевидно, ймовірності цих подій: $P(A) = 1/2$, $P(B) = 1/4$, $P(B|A) = 1/2$. Відповідні їм кількості інформації будуть такими: $I(A) = 1$, $I(B) = 2$, $I(B|A) = 1$. Таким чином

$$I(AB) = I(A) + I(B) - I(B|A) = 2.$$

Ми знову отримали той самий результат але у інший спосіб.

Важливою числовою характеристикою кількості інформації є інформаційна ентропія H . Вона характеризує не

окреме інформаційне повідомлення, а весь можливий ансамбль повідомлень $E_1, E_2, \dots, E_n$ яких може бути і нескінченно багато. Ці повідомлення покривають весь простір елементарних подій $\Omega = E_1 + E_2 + \dots + E_n$ і не мають спільних точок, крім меж. Якщо ймовірності реалізацій цих подій є $P(E_1), P(E_2), \dots, P(E_n)$, $P(E_1) + P(E_2) + \dots + P(E_n) = 1$, то інформаційна ентропія є математичним очікуванням кількості інформації

$$H = -P(E_1)\ln[P(E_1)] - P(E_2)\ln[P(E_2)] - \dots - P(E_n)\ln[P(E_n)]$$

або

$$H = -\sum_{i=1}^n P(E_i)\ln[P(E_i)].$$

Ця функція характеризує невизначеність інформації, представленої ансамблем повідомлень. Вона дорівнює нулю, коли ймовірності всіх повідомлення ансамблю, крім одного, дорівнюють нулю, а це єдине має ймовірність одиницю. Дійсно, довільний доданок останньої суми

$$\lim_{P(E_i) \rightarrow 0} H = -\lim_{P(E_i) \rightarrow 0} P(E_i)\log[P(E_i)] = 0,$$

$$\lim_{P(E_i) \rightarrow 1} H = -\lim_{P(E_i) \rightarrow 1} P(E_i)\log[P(E_i)] = 0.$$

Аналіз показує, що максимум інформаційної ентропії досягається, коли всі повідомлення ансамблю $E_1, E_2, \dots, E_n$ рівноправні і $P(E_1) = P(E_2) = \dots = P(E_n) = 1/n$. Це означає, що інформація про елементи ансамблю повідомлень максимально невизначена. Кожна з подій ансамблю подій відбуватиметься з тою самою ймовірністю.

У разі задання інформації як неперервної функції в просторі елементарних подій ми повинні пов'язати її з густиною розподілу ймовірностей повідомлень $P_\xi(x)$ і розглядати як густину інформації

$$i_\xi(x) = -\ln[P_\xi(x)].$$

При цьому інформаційна ентропія визначатиметься так

$$H = - \int_{-\infty}^{\infty} P_{\xi}(x) \ln[P_{\xi}(x)] dx.$$

Нехай один ансамбль подій, заданий в просторі елементарних подій Ω_1 , а інший в просторі елементарних подій Ω_2 . Ці простори можуть бути як залежними, так і незалежними. Тоді, для незалежних просторів

$$H(\Omega_1 \Omega_2) = H(\Omega_1) + H(\Omega_2).$$

Доведення. За означенням

$$H(\Omega_1 \Omega_2) = - \sum_{i=1}^n \sum_{j=1}^m P(E_i \bar{E}_j) \ln[P(E_i \bar{E}_j)],$$

де $\Omega_2 = \bar{E}_1 + \bar{E}_2 + \dots + \bar{E}_m$. Для незалежних подій $P(E_i \bar{E}_j) = P(E_i)P(\bar{E}_j)$. Тоді

$$\begin{aligned} H(\Omega_1 \Omega_2) &= - \sum_{i=1}^n \sum_{j=1}^m P(E_i)P(\bar{E}_j) \ln[P(E_i)P(\bar{E}_j)] = \\ &= - \sum_{i=1}^n \sum_{j=1}^m P(E_i)P(\bar{E}_j) \ln[P(E_i)] - \\ &\quad - \sum_{i=1}^n \sum_{j=1}^m P(E_i)P(\bar{E}_j) \ln[P(\bar{E}_j)]. \end{aligned}$$

Оскільки $\sum_{j=1}^m P(\bar{E}_j) = \sum_{i=1}^n P(E_i) = 1$, то

$$\begin{aligned} H(\Omega_1 \Omega_2) &= - \sum_{i=1}^n P(E_i) \ln[P(E_i)] - \sum_{j=1}^m P(\bar{E}_j) \ln[P(\bar{E}_j)] = \\ &= H(\Omega_1) + H(\Omega_2). \end{aligned}$$

Для залежних просторів елементарних подій

$$H(\Omega_1 \Omega_2) = H(\Omega_1) + H(\Omega_2 | \Omega_1),$$

де

$$H(\Omega_2 | \Omega_1) = \sum_{i=1}^n P(E_i) H(\Omega_2 | E_i)$$

умовна інформаційна ентропія ансамблю елементарних подій Ω_2 при умові що відбулись події ансамблю елементарних подій Ω_1 , $\Omega_1 = E_1 + E_2 + \dots + E_n$. Тут умовна інформаційна ентропія ансамблю елементарних подій Ω_2 при умові E_i .

$$H(\Omega_2 | E_i) = - \sum_{j=1}^n P(\bar{E}_j | E_i) \ln[P(\bar{E}_j | E_i)].$$

Доведення. За означенням

$$H(\Omega_1 \Omega_2) = - \sum_{i=1, j=1}^n P(E_i \bar{E}_j) \ln[P(E_i \bar{E}_j)],$$

Для залежних подій $P(E_i \bar{E}_j) = P(E_i)P(\bar{E}_j | E_i)$. Тоді

$$\begin{aligned} H(\Omega_1 \Omega_2) &= - \sum_{i=1, j=1}^n P(E_i)P(\bar{E}_j | E_i) \ln[P(E_i)P(\bar{E}_j | E_i)] = \\ &= - \sum_{i=1, j=1}^n P(E_i)P(\bar{E}_j | E_i) \ln[P(E_i)] - \\ &\quad - \sum_{i=1, j=1}^n P(E_i)P(\bar{E}_j | E_i) \ln[P(\bar{E}_j | E_i)]. \end{aligned}$$

Оскільки $\sum_{j=1}^m P(\bar{E}_j | E_i) = 1$, то

$$\begin{aligned} H(\Omega_1 \Omega_2) &= - \sum_{i=1}^n P(E_i) \ln[P(E_i)] - \\ &\quad - \sum_{i=1, j=1}^n P(E_i)P(\bar{E}_j | E_i) \ln[P(\bar{E}_j | E_i)] = \\ &= H(\Omega_1) + H(\Omega_2 | \Omega_1), \end{aligned}$$

де

$$H(\Omega_2 | \Omega_1) = \sum_{i=1}^n P(E_i) H(\Omega_2 | E_i),$$

$$H(\Omega_2 | E_i) = - \sum_{j=1}^n P(\bar{E}_j | E_i) \ln[P(\bar{E}_j | E_i)].$$

1.9. Задачі для самостійного розв'язання

Варіант 1.

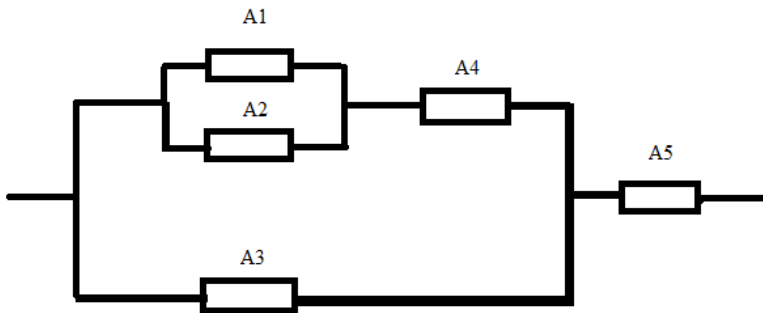
1. В колоді 36 карт. Знайти ймовірності подій:

А - чотири шістки розташовані поряд;

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

Побудувати простір елементарних подій і події А і В.

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



3. Під час випробувань було встановлено, що ймовірність безвідмовного спрацювання 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 спрацювання 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:

А - випало число, що не перевищує 2;

В - 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:

А - чотири шістки розташовані поряд;

В - місця розташування шісток утворюють геометричну 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, зсіпані 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:

A - обидва автобуси прибули своєчасно;

B - обидва автобуси запізнились;

C - 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 групу 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 урни 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 три останні цифри. Знайти ймовірності подій:

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

В - стерлись однакові цифри;

С - 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.

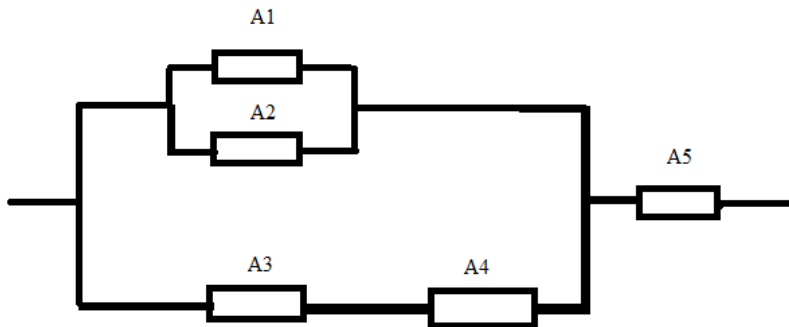
3. В будівельному загоні 70% першокурсників і 30% другокурсників. Серед першокурсників дівчата складають 10%, а серед другокурсників - 5%. Всі дівчата по черзі чергують на кухні. Яка ймовірність, що у випадково обраний день на кухні чергує першокурсниця?

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 , $k = 1, 2, \dots, 5$ за схемою, наведеною на малюнку. При виході з ладу довільного елемента коло в місці його ввімкнення розривається. Ймовірність виходу з ладу за даний період елемента A_k дорівнює p_k , $k = 1, 2, \dots, 5$. 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 $\xi = \xi(\omega)$, де $\omega \in \Omega$. 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_r - випав герб і ω_p - випала решка $\Omega = \{\omega_r, \omega_p\}$. Ми можемо, наприклад, гербу поставити у відповідність 1, а решці – 0 $1 = \xi(\omega_r)$, $0 = \xi(\omega_p)$, сформувавши випадкову величину з наступним набором можливих значень $\xi = \{0, 1\}$. Причому, це лише один з нескінченної кількості способів формування випадкової величини у даному випадку.

З максимально можливою повнотою неперервна випадкова величина описується своєю функцією розподілу

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

Тобто 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 скінченному 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 $(x_1 \leq \xi < x_2)$

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

або

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

Відповідно до аксіоми адитивності ймовірностей

$$\begin{aligned} P(\xi < x_2) &= P[(\xi < x_1) + (x_1 \leq \xi < x_2)] = \\ &= P(\xi < x_1) + P(x_1 \leq \xi < x_2). \end{aligned}$$

Мовою функцій розподілу ймовірностей останню рівність можна записати так

$$F_{\xi}(x_2) = F_{\xi}(x_1) + P(x_1 \leq \xi < x_2).$$

Звідси шукана ймовірність

$$P(x_1 \leq \xi < x_2) = F_{\xi}(x_2) - F_{\xi}(x_1).$$

Функція розподілу має наступні властивості:

1. якщо $x_1 \leq x_2$, то $F_{\xi}(x_1) \leq F_{\xi}(x_2)$. Ця властивість означає, що функція розподілу є монотонно зростаючою функцією;

$$2. \lim_{x \rightarrow -\infty} F_{\xi}(x) = F_{\xi}(-\infty) = 0; \lim_{x \rightarrow \infty} F_{\xi}(x) = F_{\xi}(\infty) = 1. \quad \text{Ця}$$

властивість є наслідком відповідної аксіоми ймовірності, про те що ймовірність неможливої події дорівнює нулю, а ймовірність достовірної події дорівнює одиниці;

3. $\lim_{x \rightarrow x_0 - 0} F_{\xi}(x) = F_{\xi}(x_0)$. 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 границі функції.

Фактично функція розподілу ймовірності є інтегральною характеристикою випадкової величини, оскільки містить всю можливу інформацію про випадкову величину на проміжку $(-\infty, x)$. 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 $P_{\xi}(x)$, яка називається густиною розподілу. Вона пов'язана з функцією розподілу або інтегральним співвідношенням

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

або диференційним співвідношенням

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

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

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

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

Густина розподілу ймовірностей має наступні властивості:

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

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

$$F_{\xi}(\infty) = \int_{-\infty}^{\infty} P_{\xi}(x) dx = 1.$$

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

$$P(a \leq x < b) = \int_a^b P_{\xi}(x) dx.$$

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 така постановка питання завжди має сенс і закон розподілу дискретної випадкової величини завжди

можна записати у вигляді таблиці з двох рядків або стовпців. Один з них - можливі значення випадкової величини

$$x_1, x_2, \dots, x_N,$$

інший – ймовірності цих значень

$$P_1, P_2, \dots, P_N,$$

де

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

Зовні закон розподілу дискретної випадкової величини має простіший вигляд.

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.

2.2. Закон розподілу кількох випадкових величин

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

$$\xi = (\xi_1, \dots, \xi_r).$$

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

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

що розглядається як функція точки $x = (x_1, \dots, x_r)$ r – вимірного евклідового простору R^r . Багатовимірною функцією розподілу $F_{\xi}(x)$ однозначно визначає ймовірність приналежності точки ξ з координатами $(\xi_1, \xi_2, \dots, \xi_r)$ довільному паралелепіпеду $B \subset R^r$,

$$B = \{(x_1, \dots, x_r) : a_1 \leq x_1 < b_1, \dots, a_r \leq x_r < b_r\}.$$

Позначимо цю ймовірність $P(\xi \in B)$. Визначимо цей зв'язок, наприклад, для двовимірного випадку, де

$$B = \{(x_1, x_2) : a_1 \leq x_1 < b_1, a_2 \leq x_2 < b_2\}.$$

У цьому разі можна показати, що

$$\begin{aligned} P[(\xi_1, \xi_2) \in B] &= \\ &= F_{\xi_1, \xi_2}(a_1, b_1) + F_{\xi_1, \xi_2}(a_2, b_2) - F_{\xi_1, \xi_2}(a_1, b_2) - F_{\xi_1, \xi_2}(a_2, b_1). \end{aligned}$$

За багатовимірними функціями розподілу можна визначити і одномірні функції розподілу. Безпосередньо можна перевірити, що

$$\begin{aligned} \lim_{y \rightarrow \infty} F_{\xi_1, \xi_2}(x, y) &= F_{\xi_1, \xi_2}(x, \infty) = F_{\xi_1}(x), \\ \lim_{x \rightarrow \infty} F_{\xi_1, \xi_2}(x, y) &= F_{\xi_1, \xi_2}(\infty, y) = F_{\xi_2}(y). \end{aligned}$$

Аналогічно одновимірному випадку можна ввести багатовимірну густину розподілу

$$P(\xi \in B) = \int \dots \int_B P_{\xi_1, \xi_2, \dots, \xi_r}(x_1, \dots, x_r) dx_1 \dots dx_r.$$

Останню рівність можна назвати багатовимірним законом розподілу. Зінтегрувавши багатовимірну густину розподілу за

всіма змінними крім однієї можна перейти до одновимірної густини розподілу. Наприклад у двовимірному випадку

$$P_{\xi_1}(x) = \int_{-\infty}^{\infty} P_{\xi_1 \xi_2}(x, y) dy,$$

$$P_{\xi_2}(y) = \int_{-\infty}^{\infty} P_{\xi_1 \xi_2}(x, y) dx.$$

Для дискретної багатовимірної випадкової величини закон розподілу визначається набором можливих значень випадкової величини і відповідним набором ймовірностей цих значень.

Серед ієрархії багатовимірних функцій розподілу ймовірностей багатовимірної випадкової величини найважливішою є функція розподілу розмірності, що відповідає розмірності багатовимірної випадкової величини. Зазначеним вище способом з неї можна отримати функції розподілу окремих компонент випадкової величини або довільної їх сукупності.

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

$$F_{\xi_1, \dots, \xi_r}(x_1, \dots, x_r) = F_{\xi_1}(x_1) \dots F_{\xi_r}(x_r),$$

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

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

Для дискретного закону розподілу

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

Крім випадкових величин можна розглядати і довільні функції випадкових величин, які, очевидно, також будуть

випадковими величинами. Якщо випадкові величини $\xi_1, \dots, \xi_r$ є незалежними, то довільні функції цих випадкових величин $\varphi_1(\xi_1), \dots, \varphi_r(\xi_r)$ також будуть незалежними випадковими величинами. Можна довести і загальніше твердження, що незалежними будуть і функції

$$f_1(\xi_1, \dots, \xi_r), \dots, f_r(\xi_1, \dots, \xi_r), \dots$$

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 $\eta = \varphi(\xi)$. Нехай функція розподілу випадкової величини ξ є відомою. Знайдемо функцію розподілу випадкової величини η . Очевидно, що

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

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

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

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

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

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

то багатовимірна густина розподілу вектора η визначається так

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

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

Задача 20. 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 децю 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 $A < 0$

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

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

2.3. Закон розподілу для функції неперервної випадкової величини

Нехай відома густина функції розподілу випадкової величини $\xi - f_{\xi}(x)$, а потрібно знайти густину функції розподілу випадкової величини $\eta - g_{\eta}(h)$. При цьому $\eta = \eta(\xi)$, $h = h(x)$.

За означенням густини розподілу ймовірності

$$dP_{\xi}(x) = P(x \leq \xi \leq x + dx) = f_{\xi}(x)dx,$$

$$dP_{\eta}(h) = P(h \leq \eta \leq h + dh) = g_{\eta}(h)dh.$$

Оскільки $\eta = \eta(\xi)$ і $h = h(x)$, то

$$dP_{\xi}(x) = dP_{\eta}(h),$$

або

$$g_{\eta}(h)dh = f_{\xi}(x)dx.$$

Звідси

$$g_{\eta}(h) = f_{\xi}(x) \frac{dx}{dh},$$

де $x = x(h)$ – обернена функція щодо функції $h = h(x)$. Через те, що похідна dx/dh може мати довільний знак, а густина

ймовірності суттєво невід’ємна функція в усіх точках своєї області визначення, то замість останнього виразу буде наступний

$$g_{\eta}(h) = f_{\xi}[x(h)] \left| \frac{dx(h)}{dh} \right|.$$

Отримана формула безпосередньо застосовна у разі, якщо обернена функція $x = x(h)$ є однозначною. Якщо ж вона m - значна, тобто може бути представлена сукупністю m - однозначних гілок $x_i = x_i(h)$, де $i = 1, \dots, m$, то отримана формула вірна для кожної окремої гілки, відповідно

$$g_{\eta}(h) = \sum_{i=1}^m f_{\xi}[x_i(h)] \left| \frac{dx_i(h)}{dh} \right|.$$

Приклад 1. Нехай $h(x) = \cos(x)$, а густина розподілу ймовірності

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

де $a = 1/\pi$, $b = 4/\pi^2$. Очевидно головне значення оберненої функції $x = \omega \cos(h)$ є функцією однозначною. Отже,

$$\begin{aligned} g_{\eta}(h) &= [a + b\omega \cos(h)] \left| \frac{d\omega \cos(h)}{dh} \right| = \\ &= [a + b\omega \cos(h)] \frac{1}{\sqrt{1-h^2}}. \end{aligned}$$

Приклад 2. Нехай $h(x) = x^2$, а густина розподілу ймовірності

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

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

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

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

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

де $\alpha = -1/2$, $\beta = 2$.

2.4. Математичне очікування

Максимально повна інформація про випадкову величину міститься в її функції розподілу. Однак на практиці часто цілком достатньою і меншою інформацією 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 $x_1, x_2, \dots, x_n$, що реалізуються з відповідними ймовірностями $P_1, P_2, \dots, P_n$, математичне очікування визначається так

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

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

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

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

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

$$P_1 = P_2 = \dots = P_n = \frac{1}{N}.$$

У цьому разі

$$M\xi = \bar{\xi}.$$

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

Для неперервної випадкової величини математичне очікування визначається через густину розподілу так

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

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

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

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

Задача 21. 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_r і ω_p : $\Omega = \{\omega_r, \omega_p\}$. Розглянемо наступну випадкову величину $\xi(\omega)$, яка приймає наступні значення

$$x_1 = \xi(\omega_r) = 1,$$

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

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

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

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

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

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

Задача 22. Обчислимо математичне очікування неперервної випадкової величини з нормальним законом розподілу

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

Розв'язання. За означенням

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

Виконаємо підстановку $x-a=y$, тоді

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

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

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

Отже математичне очікування є одним з параметрів нормального розподілу

$$M\xi = a.$$

Задача 23. Обчислимо математичне очікування неперервної випадкової величини з показниковим законом розподілу

$$P_{\xi}(x) = \alpha \exp(-\alpha x), x \geq 0.$$

Розв'язання. За означенням

$$M\xi = \int_0^{\infty} x P_{\xi}(x) dx = \alpha \int_0^{\infty} x \exp(-\alpha x) dx.$$

Інтегруючи частинами маємо

$$M\xi = -x \exp(-\alpha x) \Big|_0^{\infty} + \int_0^{\infty} \exp(-\alpha x) dx = \frac{1}{\alpha}.$$

Задача 24. Обчислимо математичне очікування дискретної випадкової величини з біноміальним розподілом

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

Розв'язання. За означенням

$$\begin{aligned} M\xi &= \sum_{m=0}^n m P_n(m) = \sum_{m=0}^n m C_n^m p^m (1-p)^{n-m} = \\ &= \sum_{m=0}^n m \frac{n(n-1)\dots(n-m+1)}{m!} p^m (1-p)^{n-m} = \end{aligned}$$

$$\begin{aligned}
&= pn \sum_{m=1}^n \frac{(n-1) \dots (n-m+1)}{(m-1)!} p^{m-1} (1-p)^{n-1-(m-1)} = \\
&= pn \sum_{k=0}^{n-1} \frac{(n-1) \dots (n-1-k+1)}{k!} p^k (1-p)^{n-1-k} = p \cdot n.
\end{aligned}$$

Оскільки умова нормування

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

Задача 25. Обчислимо математичне очікування дискретної випадкової величини з Пуассонівським розподілом

Розв'язання. За означенням

$$\begin{aligned}
M\xi &= \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} \exp(-\lambda) = \lambda \exp(-\lambda) \sum_{m=0}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \\
&= \lambda \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.
\end{aligned}$$

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 $MC\xi = CM\xi$.
3. Для довільної випадкової величини ξ

$$|M\xi| \leq M|\xi|.$$

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

4. Для довільних випадкових величин ξ_1 і ξ_2

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

5. Для незалежних випадкових величин ξ_1 і ξ_2

$$M\xi_1 \cdot M\xi_2 = M\xi_1 M\xi_2.$$

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

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

7. $P(\xi \geq \varepsilon) \leq M\xi / \varepsilon$, де випадкова величина $\xi \geq 0$ 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 $\xi(\omega) \geq \varepsilon$. 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 $\xi \geq \eta$, а отже і властивість $M\xi \geq M\eta$. Друга – Математичне очікування цієї випадкової величини легко обчислюється

$$\begin{aligned} M\eta &= \int_{\Omega} x(\omega) P_{\eta}[x(\omega)] dx(\omega) = \\ &= \int_{\Omega_\varepsilon} x P_{\eta}(x) dx = \varepsilon \int_{\Omega_\varepsilon} P_{\eta}(x) dx = \varepsilon P(\xi \geq \eta). \end{aligned}$$

Отже,

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

або

$$P(\xi \geq \varepsilon) \leq \frac{M\xi}{\varepsilon}.$$

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

Задача 26. Нехай випадкова величина ξ має нормальний розподіл з параметрами (a, σ^2) . Знайдемо математичне очікування величини $\eta = A\xi + B$.

Розв'язання. Очевидно

$$M\eta = M(A\xi + B) = AM\xi + B = Aa + B,$$

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

Задача 27. Нехай неперервна випадкова величина має рівномірний розподіл

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

Знайдемо її математичне очікування.

Розв'язання. За означенням

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

2.5. Дисперсія

Математичне очікування випадкової величини містить досить суттєву, але не вичерпну 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

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

Дійсно,

$$\begin{aligned} M(\xi - M\xi)^2 &= M[\xi^2 - 2\xi M\xi + (M\xi)^2] = \\ &= M\xi^2 - 2M(\xi M\xi) + M(M\xi)^2 = \\ &= M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2 = M\xi^2 - (M\xi)^2. \end{aligned}$$

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

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

5. $P(|\xi - M\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2}$, 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 28. Нехай неперервна випадкова величина має нормальний розподіл. Знайдемо її дисперсію.

Розв'язання. За означенням

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

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

Для обчислення використаємо інтеграл Пуассона

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

Диференціюємо його за параметром α

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

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 29. Нехай неперервна випадкова величина розподілена рівномірно. Знайдемо її дисперсію.

Розв'язання. За означенням

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

Оскільки, як було показано вище,

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

то

$$\begin{aligned} D\xi &= \int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 P_{\xi}(x) = \frac{1}{b-a} \int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 dx = \\ &= \frac{1}{b-a} \cdot \frac{1}{3} \left(x - \frac{a+b}{2}\right)^3 \Big|_a^b = \frac{(b-a)^2}{12}. \end{aligned}$$

Задача 30. Нехай дискретна випадкова величина розподілена за законом Пуассона. Знайдемо її дисперсію.

Розв'язання. За означенням

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

Останній вираз зручно записати так

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

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

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

Обчислимо перший доданок у правій частині

$$\begin{aligned} M\xi(\xi-1) &= \sum_{m=0}^{\infty} m(m-1) \frac{\lambda^m}{m!} \exp(-\lambda) = \\ &= \lambda^2 \sum_{m=2}^{\infty} \frac{\lambda^{m-2}}{(m-2)!} \exp(-\lambda) = \lambda^2 \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda^2, \end{aligned}$$

оскільки

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

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 збігання математичного очікування і дисперсії

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

Задача 31. Нехай дискретна випадкова величина розподілена за біноміальним законом. Знайдемо її дисперсію.

Розв'язання. За означенням

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

Останній вираз зручно записати так

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

Як було показано вище $M\xi = np$, тоді

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

Обчислимо перший доданок у правій частині

$$\begin{aligned} M\xi(\xi - 1) &= \sum_{m=0}^n m(m-1) \frac{n(n-1)\dots(n-m+1)}{m!} p^m (1-p)^{n-m} = \\ &= n(n-1) p^2 \sum_{m=2}^n \frac{(n-2)\dots[n-2-(m-2)+1]}{(m-2)!} p^{m-2} (1-p)^{n-2-(m-2)} = \\ &= n(n-1) p^2 \sum_{k=0}^{n-2} \frac{(n-2)\dots[n-2-k+1]}{k!} p^k (1-p)^{n-2-k} . \end{aligned}$$

Оскільки, відповідно до умови нормування

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

то

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

Остаточний результат буде таким

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

Задача 32. Розглянемо роль флуктуацій в залежності від розмірів системи У статистичній фізиці величина $D\xi$ носить назву квадратної флуктуації, оскільки характеризує відхилення випадкової величини ξ від свого середнього значення. 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 $\delta\xi \ll 1$.

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 $\delta\xi \sim 1/\sqrt{N}$.

Адитивність випадкової величини ξ означає, що її значення для цілої системи дорівнює сумі значень ξ_k $k=1, \dots, N$, де ξ_k у такій же мірі характеризує k -ту незалежну систему як ξ цілу систему, отже

$$\xi = \sum_{k=1}^N \xi_k.$$

Очевидно подібна рівність вірна 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

$$M \xi_1 = M \xi_2 = \dots = M \xi_N = \mu ,$$

$$D \xi_1 = D \xi_2 = \dots = D \xi_N = \sigma^2 ,$$

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 контроль країни окремих адміністративних одиниць, але $1/\sqrt{N}$ спадає реальна загроза від такого бунту. За малої кількості адміністративних одиниць ймовірність виходу з під контроль мала, але наслідки такої події руйнівні. Як і завжди – проблема у пошуку золотой 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 багатой людини серед громадян країни зростає пропорційно кореню квадратному з кількості громадян, а вплив такої багатой людини на перебіг подій у країні спадає обережно пропорційно тій самій величині.

Задача 33. *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 і x_2 грошей. Після зустрічі x_3 і x_4 відповідно. Густина функції розподілу двох випадкових величин ξ_1 і ξ_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 диференціальним 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 здиференціюємо спочатку за змінною 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

$$\frac{1}{P_{\xi_1}(x_1)} \frac{dP_{\xi_1}(x_1)}{dx_1} = \frac{1}{P_{\xi_2}(x_2)} \frac{dP_{\xi_2}(x_2)}{dx_2} = -\alpha,$$

де α - довільна стала. Останню подвійну рівність можна також записати так

$$\frac{1}{P_{\xi}(x)} \frac{dP_{\xi}(x)}{dx} = -\alpha,$$

де $x = x_1, x_2, \dots$. Отже отримуємо наступне звичайне диференціальне 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 формування, а саме:

$$A \int_0^{\infty} \exp(-\alpha x) dx = 1.$$

Звідси $A = \alpha$.

Таким чином, шукана густина розподілу має вигляд

$$P_\xi(x) = \alpha \exp(-\alpha x)$$

і залежить від єдиного параметра α . Для його знаходження потрібно врахувати всю суму грошей розподілених між людьми. Для цього обчислимо математичне очікування випадкової величини ξ

$$M\xi = \int_0^{\infty} x P_\xi(x) dx = \frac{1}{\alpha} = \mu.$$

З іншого боку, відповідно до закону великих чисел.

$$\lim_{N \rightarrow \infty} \bar{\xi} = \mu.$$

Тут вибіркове середнє

$$\bar{\xi} = \frac{1}{N} \sum_{k=1}^N \xi_k = \frac{E}{N}$$

де E – загальна сума грошей розподілених між людьми. Отже, якщо кількість людей достатньо велика, то

$$\bar{\xi} = \mu,$$

або

$$\mu \approx E / N$$

i

$$P_{\xi}(x) \approx \frac{N}{E} \exp(-Nx/E)$$

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 ($x \gg E/N$). Цей результат є універсальним і не залежить від якостей людей. Час його досягнення визначається лише інтенсивністю міжлюдських стосунків.

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 $w(t) = P(0 \leq \xi \leq t)$ 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

$$W(dt) = 1 - a dt .$$

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 $W(dt) \leq 1$, то $W'(0) = -a \leq 0$ для $a \geq 0$. 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 звичайне диференціальне 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

$$\tau = M\xi = \int_0^{\infty} t P_{\xi}(t) dt$$

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,

$$P_{\xi}(t) = \tau^{-1} \exp(-t / \tau)$$

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.

2.6. Коваріація. Коефіцієнт кореляції

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 характеристика, що називається коваріацією двох випадкових величин. Ця

характеристика є кількісною характеристикою залежності або незалежності випадкових величин ξ_1 і ξ_2 . За означенням

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

Використовуючи властивості математичного очікування легко отримати наступну формулу

$$\text{cov}(\xi_1, \xi_2) = M\xi_1\xi_2 - M\xi_1M\xi_2.$$

Очевидно, що якщо випадкові величини незалежні, тобто $M\xi_1\xi_2 = M\xi_1M\xi_2$, то $\text{cov}(\xi_1, \xi_2) = 0$. Якщо ж вони максимально залежні, що можливо лише у разі $\xi_1 = \xi_2$, то коваріація приймає максимально можливе значення $\text{cov}(\xi_1, \xi_2) = D\xi$. Зазначаючи властивості 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

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

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

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

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

де C – довільна стала.

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

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

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

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

Можна довести, що існує наступна оцінка коваріації

$$|\text{cov}(\xi_1, \xi_2)| \leq \sqrt{D_{\xi_1} D_{\xi_2}}.$$

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

$$\rho(\xi_1, \xi_2) = \frac{\text{cov}(\xi_1, \xi_2)}{\sqrt{D_{\xi_1} D_{\xi_2}}}.$$

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

$$|\rho(\xi_1, \xi_2)| \leq 1.$$

При цьому, як і коваріація, коефіцієнт кореляцій двох незалежних випадкових величин $\rho(\xi_1, \xi_2) = 0$. 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 $|\rho(\xi_1, \xi_2)| = 1$.

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

2.7. 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 $M(\xi - M\xi)^k$. 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 найуважнішими характеристиками випадкової величини є саме її математичне очікування та дисперсія.

Якщо ми маємо декілька випадкових величин $\xi_1, \xi_2, \dots, \xi_n$, 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

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

Центральні мішані моменти визначаються так

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

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

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

Задача 34. 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, $|\rho(\xi_1, \xi_2)| = 1$.

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

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

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

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

де ми використали властивості дисперсії, то

$$\begin{aligned} \text{cov}(\xi_1, \xi_2) &= M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)] = M[(\xi_1 - a)(\xi_2 - Aa - B)] = \\ &= M[(\xi_1 - a)(A\xi_1 + B - Aa - B)] = M[(\xi_1 - a)A(\xi_1 - a)] = \\ &= AM(\xi_1 - a)^2 = AD\xi_1 = A\sigma^2. \end{aligned}$$

Отже

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

Таким чином

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

Задача 35. Довести, що для двох незалежних випадкових величин ξ_1 і ξ_2 $\rho(\xi_1, \xi_2) = 0$.

Розв'язання. За означенням

$$\text{cov}(\xi_1, \xi_2) = M\xi_1\xi_2 - M\xi_1 M\xi_2 = M\xi_1 M\xi_2 - M\xi_1 M\xi_2 = 0.$$

3.4. 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

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

то для твірної функції моментів отримаємо аналогічне розвинення

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

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

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

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

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

то для твірної функції моментів отримаємо аналогічне розвинення

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

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

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

Отримаємо диференціальний 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

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

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

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

.....

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. Класичний ідеальний 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

$$\varepsilon_k = \frac{\hbar^2 k^2}{2m},$$

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 ймовірного знаходження електрона у стані з хвильовим вектором $\mathbf{k}$. Оскільки у кожному стані може знаходитись лише один електрон, то за означенням математичного очікування середня кількість електронів у системі, що для замкненої системи просто збігається з кількістю електронів, визначається так

$$\sum_{\mathbf{k}} 1 \cdot n(\mathbf{k}) = N_e$$

Розглянемо класичний електронний газ. З загальних принципів статистичної фізики випливає, що закон розподілу ймовірності – закон Больцмана має вигляд

$$n(k) = C \exp\left(-\frac{\varepsilon_k}{k_B T}\right).$$

Сталу величину C можна визначити з означення математичного очікування, що, фактично, є умовою нормування. Перейдемо у цій умові від сумування до інтегрування

$$\sum_{\mathbf{k}} n(k) = \frac{2V}{(2\pi)^3} \int_0^\infty dk k^2 n(k) \int_{\Omega} d\Omega$$

Оскільки закон розподілу не залежить від кутів, то інтеграл за тілесним кутом

$$\int_{\Omega} d\Omega = 4\pi$$

Тоді

$$\sum_{\mathbf{k}} n(k) = \frac{2V}{(2\pi)^3} 4\pi C \int_0^\infty dk k^2 \exp\left(-\frac{\hbar^2 k^2}{2m k_B T}\right)$$

Для обчислення інтегралу використаємо відомий результат для інтегралу Пуассона

$$\int_0^\infty dx \exp(-\alpha x^2) = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}$$

Здиференціюємо обидві частини цього виразу за α . Матимемо

$$\int_0^{\infty} dx x^2 \exp(-\alpha x^2) = \frac{1}{4} \frac{\pi^{1/2}}{\alpha^{3/2}}$$

Якщо покласти $\alpha = \frac{\hbar^2}{2mk_B T}$, то

$$\int_0^{\infty} dk k^2 \exp\left(-\frac{\hbar^2 k^2}{2mk_B T}\right) = \frac{1}{4} \frac{\pi^{1/2} (2mk_B T)^{3/2}}{\hbar^3}$$

Таким чином

$$\frac{2V}{(2\pi)^3} 4\pi C \cdot \frac{1}{4} \frac{\pi^{1/2} (2mk_B T)^{3/2}}{\hbar^3} = N$$

Звідки

$$C = 4\pi^3 \hbar^3 n \frac{1}{(2\pi mk_B T)^{3/2}}$$

Закон розподілу матиме вигляд

$$n(k) = \frac{4\pi^3 \hbar^3 n}{(2\pi mk_B T)^{3/2}} \exp\left(-\frac{\hbar^2 k^2}{2mk_B T}\right)$$

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 закону розподілу $n(k)$.

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 $[\mathbf{v}, \mathbf{v} + d\mathbf{v}]$. З загальних принципів статистичної фізики випливає, що густина розподілу

$$f(\mathbf{v}) = C \exp\left[-\frac{\varepsilon_k(\mathbf{v})}{k_B T}\right],$$

де кінетична енергія

$$\varepsilon(\mathbf{v}) = \frac{mv^2}{2}$$

Стала C 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

$$f(\mathbf{v}) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} \exp\left(-\frac{mv^2}{2k_B T}\right)$$

Густина цього розподілу називається густиною розподілу Максвелла. 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

$$\langle v_x \rangle = \int v_x f(\mathbf{v}) d\mathbf{v} = \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_z v_x f(\mathbf{v})$$

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

$$\begin{aligned} \langle v \rangle &= 2\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \int_0^\infty dx x \exp\left(-\frac{mx}{2k_B T}\right) = \\ &= 2\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \left\{ -\frac{2k_B T}{m} x \exp\left(-\frac{mx}{2k_B T}\right) \Big|_0^\infty + \frac{2k_B T}{m} \exp\left(-\frac{mx}{2k_B T}\right) \int_0^\infty \exp\left(-\frac{mx}{2k_B T}\right) dx \right\} = \\ &= 2\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \left(\frac{2k_B T}{m} \right)^2. \end{aligned}$$

Остаточно

$$\langle v \rangle = \left(\frac{2k_B T}{m} \right)^{1/2} \frac{2}{\sqrt{\pi}}, \quad 2/\sqrt{\pi} \approx 1.13.$$

Середня кінетична енергія 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

$$\int_0^\infty dx \exp(-\alpha x^2) = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}$$

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

$$\langle \varepsilon_k \rangle = 2\pi m \left(\frac{m}{2\pi k_B T} \right)^{3/2} \frac{3\pi^{1/2}}{8} \left(\frac{2k_B T}{m} \right)^{5/2}.$$

Середня квадратична швидкість

$$\sqrt{\langle v^2 \rangle} = \sqrt{\frac{2}{m} \varepsilon_{kin}} = \sqrt{\frac{3k_B T}{m}}$$

Очевидно

$$\sqrt{\langle v^2 \rangle} = \sqrt{\frac{3\pi}{8}} v.$$

Дисперсія швидкості електрона. За означенням дисперсії

$$\langle [v - \langle v \rangle]^2 \rangle = \langle v^2 \rangle - \langle v \rangle^2$$

Оскільки раніше нами були отримані наступні результати:

$$\langle v \rangle = \left(\frac{8k_B T}{\pi m} \right)^{1/2},$$

$$\langle \varepsilon_k \rangle = \frac{m}{2} \langle v^2 \rangle = \frac{3}{2} k_B T,$$

то

$$\begin{aligned}\langle v^2 \rangle &= \frac{3k_B T}{m} \\ \langle [v - \langle v \rangle]^2 \rangle &= \frac{3k_B T}{m} - \frac{8k_B T}{m\pi} = \frac{k_B T}{m} \left(3 - \frac{8}{\pi} \right) \\ \langle [v - \langle v \rangle]^2 \rangle &= \frac{k_B T}{m} \left(3 - \frac{8}{\pi} \right)\end{aligned}$$

Дисперсія кінетичної енергії електрона.

$$\begin{aligned}\langle [\varepsilon_{kin} - \langle \varepsilon_{kin} \rangle]^2 \rangle &= \langle \varepsilon_{kin}^2 \rangle - \langle \varepsilon_{kin} \rangle^2 = \\ &= \frac{m^2}{4} \left[\langle v^4 \rangle - \langle v^2 \rangle^2 \right].\end{aligned}$$

Для середнього квадрату швидкості раніше нами був отриманий результат

$$\langle v^2 \rangle = \frac{3k_B T}{m}.$$

Обчислимо $\langle v^4 \rangle$. За означенням

$$\begin{aligned}\langle v^4 \rangle &= \int v^4 f(v) dv = 4\pi \int_0^\infty dv v^6 f(v) = \\ &= 4\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \int_0^\infty dv v^6 \exp \left[-\frac{mv^2}{2k_B T} \right]\end{aligned}$$

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

$$\begin{aligned}\langle v^4 \rangle &= 4\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} \frac{15\pi^{1/2}}{16} \left(\frac{2k_B T}{m} \right)^{7/2} = \\ &= \frac{15\pi^{1/2}}{4} \left(\frac{2k_B T}{m} \right)^2 = 15 \left(\frac{k_B T}{m} \right)^2\end{aligned}$$

Остаточню

$$\langle [\mathcal{E}_{kin} - \langle \mathcal{E}_{kin} \rangle]^2 \rangle = \frac{m^2}{4} \left\{ 15 \left(\frac{k_B T}{m} \right)^2 - 9 \left(\frac{k_B T}{m} \right)^2 \right\} = \frac{3}{2} (k_B T)^2$$

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

$$dn(v_x) = n f(v_x) dv_x$$

- кількість молекул одиниці об'єму, компонента швидкості v_x яких в інтервалі $[v_x, v_x + dv_x]$. 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

$$\int_0^{\infty} dx x^2 \exp(-\alpha x^2) = \frac{\pi^{1/2}}{4} \alpha^{-3/2},$$

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

$$P = nk_B T$$

2.9. Закон великих чисел

Закон великих чисел описує поведінку середнього арифметичного n випадкових величин при $n \rightarrow \infty$. 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 $\xi_1, \xi_2, \dots, \xi_n$. Фактично закон великих чисел моделює ситуацію з вимірюванням тої чи іншої величини ξ , де $\xi_1, \xi_2, \dots, \xi_n$ 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 границі $n \rightarrow \infty$, 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 $\xi_1, \xi_2, \dots, \xi_n$ однаково розподілені,

попарно незалежні і мають скінчені дисперсії, тоді для довільного $\varepsilon > 0$

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

Тут

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

– середнє арифметичне всіх випадкових величин, $M_{\xi_k} = a$.

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

$$M_{\xi} = \lim_{n \rightarrow \infty} \bar{\xi}.$$

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

$$\lim_{n \rightarrow \infty} P \left\{ \left| \bar{\xi} - \overline{M_{\xi}} \right| < \varepsilon \right\} = 1,$$

де $\bar{\xi}$ – 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 $n \rightarrow \infty$. 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

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

Використаємо цю нерівність для оцінки наближеного значення вимірюваної величини. Нехай виконане n незалежних вимірювань певної невідомої величини a . Похибки вимірювань $\delta_1, \delta_2, \dots, \delta_n$ будемо вважати випадковими величинами. Припустимо, що $M\delta_k = 0$, $1 \leq k \leq n$. 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:

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

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,

$$P(|\eta| \geq \varepsilon) \leq \frac{D\eta}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Якщо

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

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:

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

2.10. Граничні теореми в схемі Бернуллі

Як відомо, послідовність n незалежних випробувань з двома наслідками описується біноміальним розподілом або розподілом Бернуллі

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

Тут m – кількість успіхів в n – випробуваннях, $n-m$ – кількість невдач, p – ймовірність успіху, $q=1-p$ – ймовірність невдачі. В міру зростання кількості випробувань технічне використання формули Бернуллі ускладнюється. Однак виявляється, що при досить великій кількості випробувань замість точної формули Бернуллі можна використовувати наближене граничне її значення. Причому відміна між точною і наближеною формулами для достатньо великих n може стати як завгодно малою. Принциповою для математичної статистики є та обставина, що за досить загальних умов, граничною функцією розподілу є нормальна функція розподілу. Останнє свідчить про те, що центральну роль у математичній статистиці грає саме нормальний закон розподілу випадкових величин. Це твердження і становить зміст центральної граничної теореми.

Наведемо формулювання декількох основних граничних теорем про які вже побіжно згадувалось вище.

Теорема Пуассона. Якщо $n \rightarrow \infty$ і $p \rightarrow 0$ так, що $n \cdot p \rightarrow \lambda$, $0 < \lambda < \infty$, то

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

Таким чином, для великих n і малих p ми можемо наближено скористатись формулою

$$P_n(m) \approx \frac{\lambda^m}{m!} \exp(-\lambda), \quad \lambda = n \cdot p.$$

Наведемо декілька чисельних прикладів, що показують степінь точності граничної формули Пуассона. Нехай $m=0$, $n=10$, $p=1/5$ і $p=1/50$ для $n=100$, тоді

$$P_{m=0}(\lambda=2) = 0.1353,$$

$$P_{n=10}(m=0) = 0.1074,$$

$$P_{n=100}(m=0) = 0.1326.$$

У формулу Пуассона n не входить, а формула Бернуллі для малих n суттєво залежить від n . З наведеного прикладу видно, що для $n=10$ розбіжність між формулами Бернуллі і Пуассона виникає у другій значущій цифрі, а при $n=100$ лише у третій. Отже, для $m=0$ великими значеннями $n \in n \approx 100$. Аналогічний вигляд мають дані і для $m=1$, $n=10$ $p=1/5$ і $p=1/50$ для $n=100$:

$$P_{m=1}(\lambda=2) = 0.2707,$$

$$P_{n=10}(m=1) = 0.2684,$$

$$P_{n=100}(m=1) = 0.2707.$$

Якщо малою є ймовірність невдачі $q=1-p$, то формулу Пуассона можна використовувати для числа невдач. Якщо ж

обидві ймовірності p і $1-p$ не є малими, то звичайно використовуються локальна і глобальна теореми Муавра–Лапласа.

Локальна теорема Муавра–Лапласа. Якщо $n \rightarrow \infty$, $p(0 < p < 1)$ є сталою, то випадкова величина $x_m = (m - np) / \sqrt{npq}$ розподілена за нормальним законом з параметрами $(0, 1)$

$$P_n(m) = \varphi(x_m),$$

де густина нормального розподілу з математичним очікуванням 0 і математичною дисперсією 1

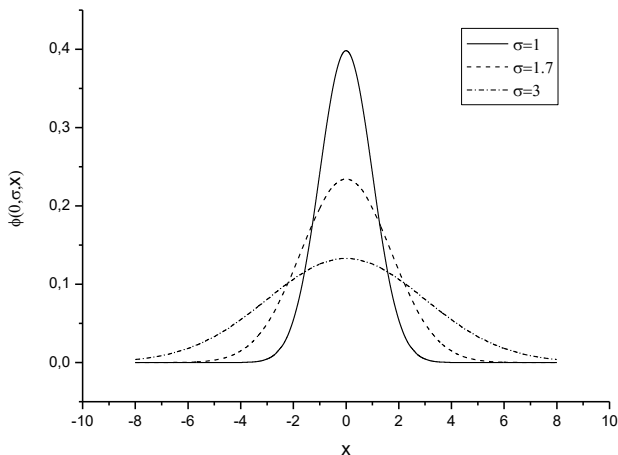
$$\varphi(x_m) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_m^2}{2}\right).$$

В загальному випадку густина нормального розподілу має вигляд

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

де μ - математичне очікування, а σ^2 - дисперсія для нормального розподілу. Зауважимо, що з умови $n \rightarrow \infty$ випливає, що і $m \rightarrow \infty$. Останнє означає, що при практичному використанні наближеної формули необхідно, щоб n і m були одного порядку величини. Оцінка точності наближеної формули є досить складною, тому цією формулою потрібно користатись дуже обережно. Принциповою обставиною є не стільки практична цінність отриманої наближеної формули, скільки теоретична, оскільки за певних умов граничною формою розподілу Бернуллі є нормальний розподіл.

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



Мал. 12. Густина нормального розподілу для різних значень дисперсії

Нормальний розподіл є межовим розподілом не тільки для розподілу Бернуллі але і для більшості інших розподілів ймовірності

Інтегральна теорема Муавра–Лапласа. При виконанні умов локальної теореми Муавра–Лапласа закон розподілу можна записати так

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

Останній вираз можна записати так

$$P(a \leq x_m \leq b) \approx \Phi(b) - \Phi(a)$$

або

$$P(a \leq x_m \leq b) \approx \Phi_0(b) - \Phi_0(a) ?$$

де

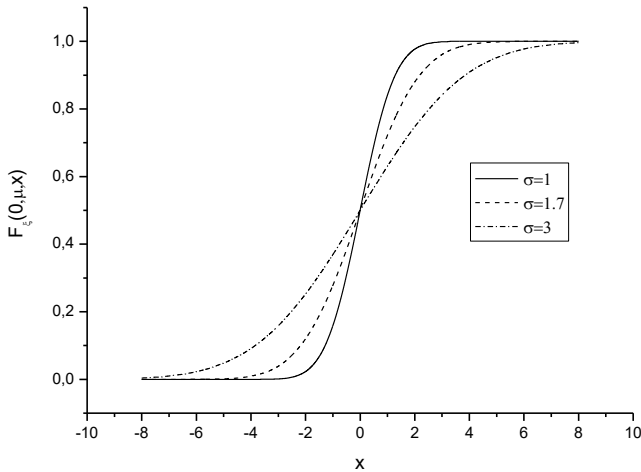
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-y^2/2) dy,$$

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

- функція Лапласа. Очевидно, що ця функція є непарною. Наведена функція $\Phi(x)$ є функцією розподілу ймовірностей випадкової величини з нульовим математичним очікуванням і одиничною дисперсією. В загальному випадку цю функцію розподілу можна записати так

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

Графік цієї функції наступний



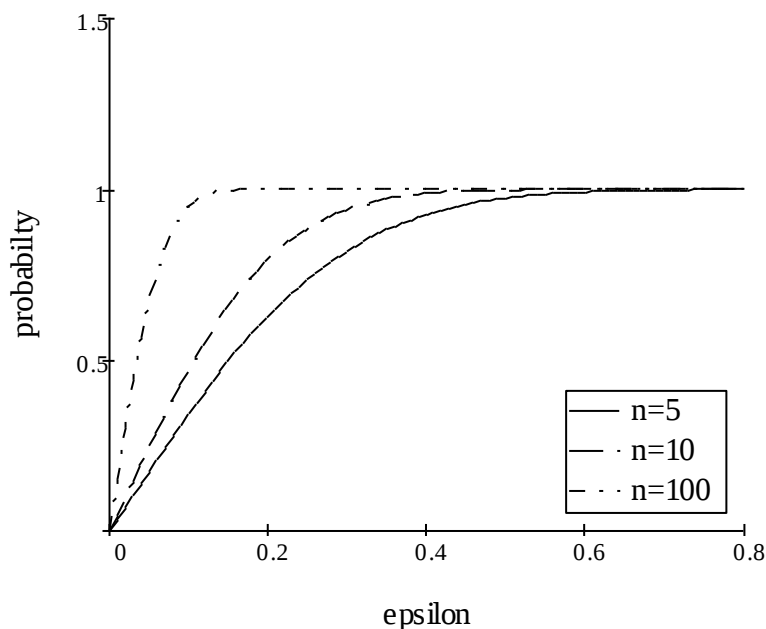
Мал. 13. Функція нормального розподілу в залежності від дисперсії

Як вже говорилося вище, інтегральна формула Муавра–Лапласа дозволяє оцінити близькість частоти успіху і його

ймовірності. Частота успіху визначається як m/n . Оцінимо ймовірність події $\left| \frac{m}{n} - p \right| < \varepsilon$. Якщо n достатньо велика, то можна скористатись інтегральною формулою Муавра–Лапласа. Тоді

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

Якщо n і p відомі, то для довільного ε ми зможемо знайти відповідну ймовірність того, що різниця між частотою і ймовірністю не виходять за межі інтервалу $(-\varepsilon, \varepsilon)$. Залежність цієї ймовірності від величини інтервалу і від кількості випробувань можна побачити на наступному малюнку.



Мал. 14. Ймовірності відхилення від розмірів інтервалу і від кількості випробувань n ($p = 0.5$).

Видно, що чим більша кількість випробувань, тим швидше ймовірність приналежності частоти певному інтервалу зростає із збільшенням величини цього інтервалу і швидше досягає максимально можливого значення 1.

З останньої формули легко знайти кількість випробувань, яку потрібно здійснити, щоб різниця між частотою і ймовірністю успіху належала певному інтервалу. Для цього потрібно розв'язати останнє рівняння щодо величини

$$u_{\alpha} = \varepsilon \sqrt{\frac{n}{p(1-p)}},$$

вважаючи ймовірність $P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right)$ відомою величиною. В які можна тут використати. Можна скористатись і будь-яким прикладним математичним пакетом типу Mathcad, де є відповідні вбудовані процедури-функції. Вважаючи, таким чином, величину u_{α} відомою для кожного значення ймовірності, матимемо для відповідної кількості випробувань

$$n \geq \frac{u_{\alpha}^2}{\varepsilon^2} p(1-p).$$

Отже, відповідне значення кількості випробувань має бути не меншим від отриманого.

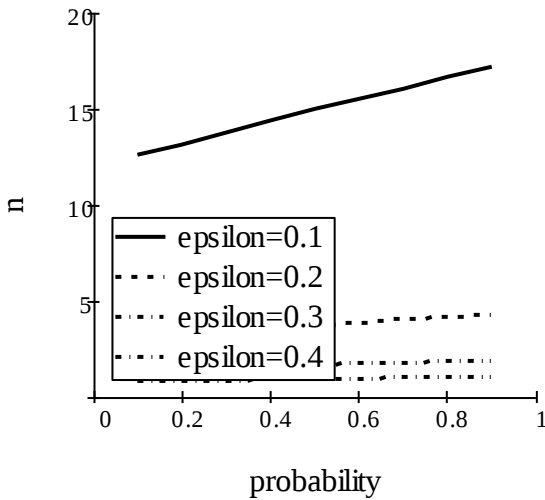
Якщо ймовірність успіху невідома, то можна використати ту обставину, що

$$p(1-p) \leq 1/4.$$

Тоді передостаннє співвідношення можна записати у вигляді нерівності

$$n \geq \frac{u_{\alpha}^2}{4\varepsilon^2}.$$

На наступному малюнку можна побачити залежність кількості випробувань від ймовірності.



Мал. 15. Залежність кількості випробувань від ймовірності приналежності частоти заданому інтервалу ($p = 0.5$).

З малюнка видно, наприклад, що для інтервалу $(-0.1, 0.1)$ кількість випробувань не перевищує 20 для, щоб частота відрізнялась від ймовірності успіху менше, ніж на 0.1.

Задача 36.. Довести, що для $n \rightarrow \infty$ і $p \rightarrow 0$ так, що $n \cdot p \rightarrow \lambda$, $0 < \lambda < \infty$, $P_n(m) = \frac{\lambda^m}{m!} \exp(-\lambda)$.

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

$$\begin{aligned} P_n(m) &= C_n^m p^m (1-p)^{n-m} = \frac{n(n-1)\dots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{n-m} = \\ &= \frac{\lambda^m}{m!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m}. \end{aligned}$$

Перейдемо тепер до границі $n \rightarrow \infty$

$$\begin{aligned}\lim_{n \rightarrow \infty} P_n(m) &= \lim_{n \rightarrow \infty} \left[\frac{\lambda_n^m}{m!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-m} \right] = \\ &= \frac{\lambda_n^m}{m!} \lim_{n \rightarrow \infty} \left[\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \right] \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^n \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^{-m} = \\ &= \frac{\lambda^m}{m!} \exp(-\lambda).\end{aligned}$$

Тут ми використали той факт, що

$$\begin{aligned}\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^n &= \exp(-\lambda), \\ \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^{-m} &= 1, \\ \lim_{n \rightarrow \infty} \left[\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \right] &= 1.\end{aligned}$$

Нормальність закону розподілу для випадкової величини $x_m = (m - np) / \sqrt{npq}$ для достатньо великих n – кількості випробувань в схемі Бернуллі є підґрунтям для центральної граничної теореми. Фактично ця теорема збігається з інтегральною або локальною теоремами Муавра–Лапласа, але записаними не для однієї випадкової величини, а для їх довільної суми. Центральна гранична теорема має декілька формулювань. Наведемо одне з них.

Теорема. Якщо випадкові величини $\zeta_1, \zeta_2, \dots, \zeta_n, \dots$ незалежні, однаково розподілені і мають скінчену дисперсію, то для великих n їх середнє арифметичне $\bar{\xi}$ розподілене за нормальним законом з параметрами $(a, \sigma^2 / n)$,

де $a = M\xi_i$, $\sigma^2 = D\xi_i$, $\bar{\xi} = (\xi_1 + \xi_2 + \dots + \xi_n) / n$.

Локальний варіант цієї теореми можна записати так

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

Відповідно, і інтегральний варіант

$$F_{\bar{\xi}}(x) = P(\bar{\xi} < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x P_{\bar{\xi}}(y) dy.$$

Закон великих чисел визначає границю, до якої прямує вибіркове середнє $\bar{\xi}$ для $n \rightarrow \infty$. Центральна гранична теорема визначає поведінку випадкової величини $\bar{\xi}$ для достатньо великих але скінчених n . Відповідно до цього закону ця поведінка виявляється універсальною (нормальний закон розподілу) незалежно від закону розподілу випадкових величин, що утворюють вибіркове середнє.

Формула Шеннона. Нехай ζ – дискретна випадкова величина, що може приймати значення $x_1, x_2, \dots, x_N$ 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 Шеннону кількість інформації в бітах визначається наступною формулою

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

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

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

і ми одержуємо формулу Хартлі (американський математик).

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 Шеннона, інформація, отримана із відповіді, є наступною

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

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 $P = q = 1/2$ ($I = \log_2 2 = 1$).

2. 11. Задачі для самостійного Розв'язання

Варіант 1.

1. На кожні 40 відштампованих виробів у середньому припадає 4 дефектних. Із усієї продукції навання 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; 2) не менше, ніж 35 разів; в) від 32 до 55 разів; г) відхилення числа появ 1-го сектора від найімовірнішого значення не перевищує 10 появ?

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 повинне бути максимальне відхилення частоти від імовірності, щоб ймовірність цього відхилення дорівнювала 0.8?

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 ймовірності не перевищувало 0.1?

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 ймовірності, щоб ймовірність цього відхилення дорівнювала 0.9?

Варіант 4.

1. Зерна пшениці проростають з ймовірністю 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?

Варіант 5.

1. Ймовірність розпаду радіоактивного атому протягом року дорівнює 0.0001. Яка ймовірність того, що з 3000 0 атомів за рік розпадутися: 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 .

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?

Варіант 7.

1. Відомо, що ймовірність влучення в ціль при одному пострілі дорівнює 0.21. Спортсмен 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 ймовірністю 0.94 відхилення частоти від імовірності не перевищувало 0.15?

Варіант 8.

1. В лотереї розігрується багато призів. Ймовірність 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?

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?

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 .

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. Результати вимірювань є, найчастіше, головним джерелом інформації про вимірювану величину. Їх характерною рисою є відмінність між собою результатів окремих вимірювань. Набір результатів окремих вимірювань найчастіше виглядає як сукупність значень певної випадкової величини. Адекватною математичною моделлю вимірюваного об'єкту є, так звана, генеральна сукупність. Під останньою розуміється ймовірнісний простір (сукупність простору елементарних подій, алгебри подій і заданих на ній ймовірностей), що є областю визначення випадкової величини ξ , яка, у свою чергу, відповідає досліджуваній характеристиці об'єкту a . Одиницею генеральної сукупності є певне значення випадкової величини або випадкового вектора. Розглянемо приклад дискретної і скінченної генеральної сукупності. Нехай досліджуються певні характеристики, наприклад середній зріст, певного народу. Тут в якості одиниці генеральної сукупності можна взяти окремого представника даного народу і значення випадкової величини, що йому відповідає - його зріст. Прикладом неперервної генеральної сукупності є набір результатів вимірювань геометричних параметрів довільного об'єкту, наприклад, довжини. Тут елементом генеральної сукупності є результат окремого вимірювання. Як правило, до початку вимірювань, у нас відсутня інформація про генеральну сукупність. Єдиним джерелом про

генеральну сукупність (ймовірнісний простір) є процедура вимірювання.

У разі дискретної скінченної генеральної сукупності теоретично існує можливість безпосередньо обчислити шукану середню характеристику даної генеральної сукупності, хоча практично це може бути недоцільним. У разі неперервної генеральної сукупності безпосереднє точне обчислення шуканої характеристики об'єкту неможливе навіть теоретично. В обох випадках виникає потреба замінити вихідну генеральну сукупність простішою математичною моделлю, в рамках якої досліджувану характеристику генеральної сукупності можна обчислити з наперед заданою точністю. Такою математичною моделлю є вибірка з генеральної сукупності. Вибірka - це довільна сукупність елементів генеральної сукупності. Наприклад, нехай потрібно виміряти певну величину a . Результати вимірювань цієї величини $x_1, x_2, \dots, x_n$ можна розглядати як набір значень певної дискретної 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 . На основі вибірки ми можемо з контрольованою точністю оцінити і саму вимірювану величину a . Це можна зробити, наприклад, взявши середнє арифметичне всіх елементів вибірки, тобто

$$a \approx \bar{\xi} = \frac{1}{n} \sum_{m=1}^n x_m.$$

Проте сприйняття елементів вибірки, як звичайних дійсних чисел, не дозволяє розбудувати серйозний теоретичний апарат дослідження вибірки, як складного математичного об'єкту. Для подальшої розбудови математичного апарату математичної статистики доцільно розглядати кожен з величин $x_1, x_2, \dots, x_n$, у свою чергу, як випадкову величину. 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

$$\begin{aligned} &X_{11}, X_{12}, \dots, X_{1n}, \\ &X_{21}, X_{22}, \dots, X_{2n}, \\ &\dots\dots\dots, \\ &X_{m1}, X_{m2}, \dots, X_{mn}, \\ &\dots\dots\dots \end{aligned}$$

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

$$\begin{aligned} \xi_1 &= \{X_{11}, X_{12}, \dots, X_{1n}\}, \\ \xi_2 &= \{X_{21}, X_{22}, \dots, X_{2n}\}, \\ &\dots\dots\dots, \\ \xi_m &= \{X_{m1}, X_{m2}, \dots, X_{mn}\}, \\ &\dots\dots\dots. \end{aligned}$$

Фактично ми замінили одну вибірку обсягом $m \times n$, що є набором можливих значень однієї випадкової величини ξ , сукупністю m вибірок обсягами по n елементів, що є наборами можливих значень m випадкових величин.

Відповідно, кожна з випадкових величин $\xi_1, \xi_2, \dots, \xi_n$, оскільки вже не є числом, має певну функцію розподілу ймовірностей. Середнє арифметичне елементів вибірки тепер також вже не буде числом, а отже до його аналізу можна застосувати весь апарат математичної статистики. Якщо всі

вимірювання проводяться у приблизно однакових умовах, то ці функції розподілу мають також бути приблизно однаковими з однаковими параметрами розподілів.

Вибіркові спостереження дають можливість зробити висновки відносно генеральної сукупності шляхом дослідження лише її скінченної частини. Для того, щоб ці спостереження - статистичні оцінки були корисними, вони мають відповідати певним вимогам. Першою такою вимогою є те, що випадкова величина $\bar{\xi}$ є незміщеною оцінкою вимірюваної величини a , тобто, що

$$M\bar{\xi} = a.$$

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

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

Із другого виразу випливає перший але не навпаки, дійсно,

$$M\bar{\xi} = M\left(\frac{1}{n} \sum_{m=1}^n \xi_m\right) = \frac{1}{n} M\left(\sum_{m=1}^n \xi_m\right) = \frac{1}{n} \sum_{m=1}^n M\xi_m = M\xi_k = a.$$

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. Оцінювання невідомого параметра a на основі результатів його вимірювання $x_1, x_2, \dots, x_n$ є однією з задач математичної статистики.

Введене вище поняття незміщеності оцінки можна дещо розширити. У певній ситуації незміщеність може бути властивістю лише вибірок достатньо великого обсягу.

Послідовність оцінок $\bar{\xi}_n$, тут

$$\overline{\xi_n} = \frac{1}{n} \sum_{m=1}^n x_m, \quad n = 1, 2, \dots$$

називається асимптотично незміщеною послідовністю оцінок параметра a , якщо

$$\lim_{n \rightarrow \infty} M(\overline{\xi_n} - a) = 0$$

або

$$\lim_{n \rightarrow \infty} M \overline{\xi_n} = a.$$

Оскільки наша оцінка параметру a є випадковою величиною, то її значення, навіть за великого обсягу вибірки, є непередбачуваними і можуть як завгодно відрізнятись від межового значення, тому важливою є наступна властивість оцінки.

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

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

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 може стосуватись не лише безпосередньо вимірюваної величини. Вона може стосуватись якогось параметру функції розподілу довільної випадкової величини. У цьому разі оцінкою вважається довільна функція елементів вибірки, а не лише використовуване нами середнє арифметичне. У цьому разі

важливою властивістю оцінки є її ефективність. 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 $F_{\xi}(x)$, або набору функцій розподілу $F_{\xi_1}(x), F_{\xi_2}(x), \dots, F_{\xi_n}(x)$ 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 $(\xi_1, \xi_2, \dots, \xi_n)$.

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

Однією з основних характеристик вибірки є емпірична функція розподілу $F_n^*(x)$, що визначається формулою

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

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

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

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}.$$

Очевидно, що

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

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

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

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

$$(-\infty, \infty) = \bigcup_{k=1}^r [z_k, z_{k+1}],$$

де

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

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

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

Останнє співвідношення нагадує відповідне співвідношення для неперервної випадкової величини

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

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

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

Оскільки при відповідному виборі точок z_k

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

то

$$P_k^* \approx P_{\xi}(x_k) \Delta z_k$$

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

$$m_v = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^v ,$$

де

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

– перший початковий вибірковий момент або вибіркове математичне очікування, або просто вибіркове середнє. Зазначимо, що вибіркове середнє виглядає як частинний випадок математичного очікування дискретної випадкової величини при умові, що ймовірність всіх значень випадкової величини однакові і дорівнюють $1/n$. Дійсно, якщо

$$M\xi = \sum_{k=1}^n p_k x_k ,$$

а $p_k = 1/n$, то

$$M\xi = \frac{1}{n} \sum_{k=1}^n x_k = \bar{x} .$$

Фактично вибіркове середнє – це математичне очікування дискретної випадкової величини, значення якої збігаються з значеннями елементів вибірки, при відсутності детальної інформації про ймовірності цих значень. Найпростішим припущенням у такій ситуації є припущення про однаковість всіх цих ймовірностей. Це припущення працює тим точніше, чим більшим є обсяг вибірки.

Для вибіркової дисперсії або другого центрального моменту випадкової величини часто використовують позначення

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

Вибіркову дисперсію також можна розглядати як дисперсію дискретної випадкової величини, значення якої збігаються із

значеннями елементів вибірки, а щодо ймовірностей цих значень робляться припущення про їх однаковість. Дійсно, оскільки

$$D\xi = M(\xi - M\xi)^2 = \sum_{k=1}^n p_k (x_k - M\xi)^2,$$

то в припущенні $p_k = 1/n$ ми отримуємо

$$D\xi = \sum_{k=1}^n p_k (x_k - M\xi)^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 = s^2.$$

Вибіркові моменти і початкові, і центральні є випадковими величинами але ступінь їх випадковості є меншим за ступінь випадковості вихідної випадкової величини і зменшується із збільшенням обсягу вибірки. Нагадаємо, що вище обговорювалась теорема про те, що дисперсія суми n однаково розподілених незалежних випадкових величин є в n разів меншою за дисперсію окремих доданків. Наше ототожнення вибірки з послідовністю однаково розподілених випадкових величин якраз і дозволяє перенести цей результат на суму елементів вибірки.

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 $F_n^*(x)$. 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 α_v , m_v бувають близькими до відповідних характеристик теоретично існуючої, але нам невідомої функції $F_\xi(x)$, тобто до її відповідних моментів

$$\alpha_v = M\xi^v$$

та

$$\mu_{\alpha} = M(\xi - M\xi)^{\nu}.$$

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

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

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

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

$$\begin{aligned} D\bar{x} &= D \frac{1}{n} \sum_{k=1}^n x_k = \frac{1}{n^2} D \sum_{k=1}^n x_k = \frac{1}{n^2} \sum_{k=1}^n Dx_k = \\ &= \frac{1}{n^2} \sum_{k=1}^n \mu_2 = \frac{\mu_2}{n}. \end{aligned}$$

Таким чином,

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

$$D\bar{x} = \frac{\mu_2}{n}$$

І ми отримали обговорюваний вище факт про те, що вибірова дисперсія для вибірки обсягом n в n разів менша за дисперсію

окремого елемента вибірки, якщо його розглядати як випадкову величину. Аналогічно можна знайти, що для $v > 1$

$$Ma_v = \alpha_v ,$$

$$Da_v = \frac{\alpha_{2v} - \alpha_v^2}{n} .$$

Складнішим є обчислення математичного очікування і дисперсії центральних вибірових 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 вибірових моментів збігається з відповідними теоретичними моментами. Дисперсія початкових вибірових 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 відмінність теоретичних і експериментальних моментів швидко прямує до нуля

$$Mm_v = \mu_v + D\left(\frac{1}{n}\right) ,$$

що добре видно на прикладі другого вибірового моменту, для якого

$$Mm_2 = \mu_2 - \frac{\mu_2}{n} .$$

3.5. Вибіркові моменти для великих вибірок

Для оцінки поведінки вибірових моментів для великих обсягів вибірки $n \rightarrow \infty$ важливими є наступні твердження.

Теорема 1. Якщо скінченими є теоретичні моменти парних порядків $\alpha_{2\nu}$, то для $n \rightarrow \infty$ вибіровий момент a_ν асимптотично

нормальний з параметрами $\left(\alpha_\nu, \frac{\alpha_{2\nu} - \alpha_\nu^2}{n} \right)$.

Теорема 2. Якщо скінченим є теоретичний момент μ_4 , то для $n \rightarrow \infty$ вибірова дисперсія m_2 асимптотично нормальна з

параметрами $\left(\mu_2, \frac{\mu_4 - \mu_2^2}{n} \right)$.

Отже, межові функції розподілу вибірових моментів є нормальними, що є прямим наслідком центральної межової теореми. Однак, domeжові функції розподілу різних вибірових характеристик можуть суттєво відрізнятись від нормальних, тому знаходження функцій розподілу для вибірок скінченного обсягу є надзвичайно складною задачею. При цьому, якщо в рамках обмежень теорем 1 і 2 межові функції розподілу вибірових моментів не залежать від функцій розподілу випадкових величин, то domeжові функції розподілу можуть залежати досить суттєво. Фактично більш–менш простою задачею є знаходження domeжових функцій розподілу лише у разі, випадкових величин $\xi_0, \xi_1, \dots, \xi_n$, що утворюють вибірку, і мають нормальний розподіл. Нагадаємо, що функції розподілу $F_\xi(x)$ або $F_{\xi_1}(x_1), F_{\xi_2}(x_2), \dots, F_{\xi_n}(x_n)$ – характеризують вимірювальний прилад. У разі нормально розподілених випадкових величин $\xi_0, \xi_1, \dots, \xi_n$ з параметрами $(0,1)$, розподіл випадкової величини

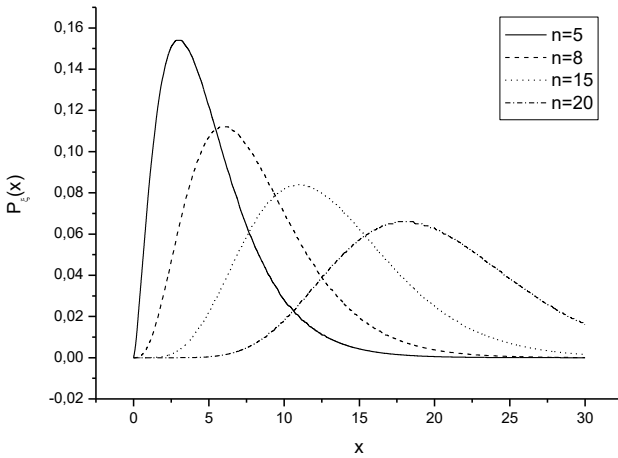
$$\chi_n^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2,$$

відомий і називається розподілом χ^2 з n – ступенями вільності.

Густина цього розподілу має вигляд

$$P_{\chi_n^2}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad x > 0, \quad \chi_n^2 = \xi_1^2 + \dots + \xi_n^2.$$

Його графік наступний

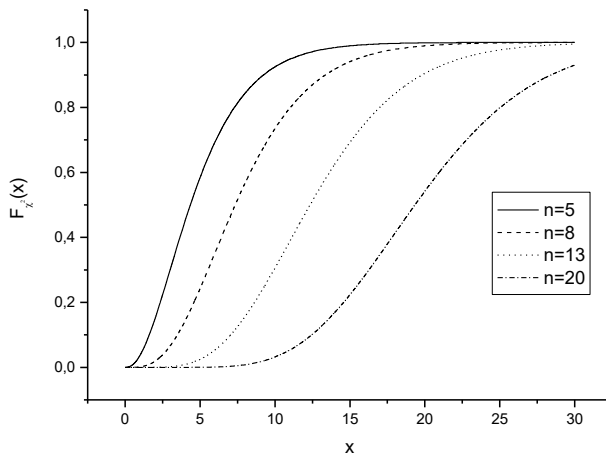


Мал. 16. Густина розподілу χ_n^2 для різних значень n .

Функція розподілу χ^2

$$F_{\chi_n^2}(x) = \int_0^x P_{\chi_n^2}(x) dx$$

має вигляд



Мал. 17. 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 Ст'юдента з n ступенями вільності.

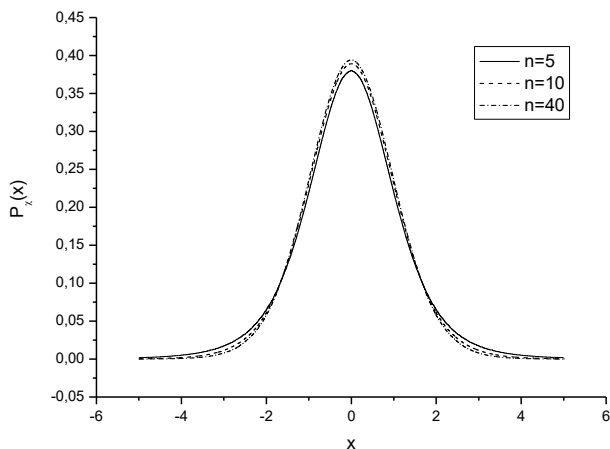
Густина розподілу Ст'юдента має вигляд

$$P_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad \tau_n = \frac{\xi_0}{\sqrt{\frac{1}{n} \chi_n^2}}, \quad x \in (-\infty, \infty).$$

Відповідно для функції розподілу матимемо

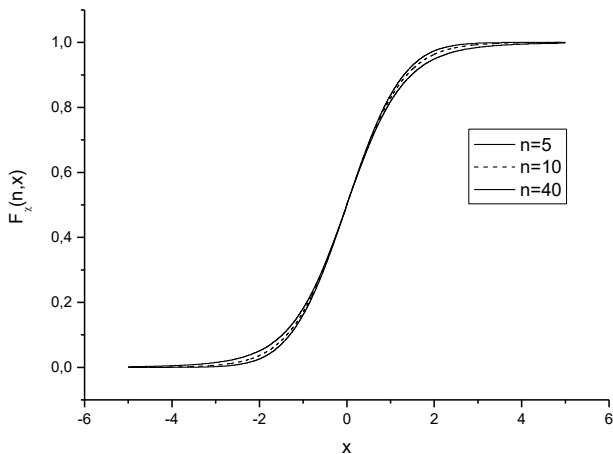
$$F_{\tau_n}(x) = \int_{-\infty}^x P_{\tau_n}(x) dx$$

Графік густини розподілу наступний



Мал. 18. Густина розподілу Ст'юдента для різних n .

Графік функції розподілу Ст'юдента наступний



Мал. 19. Функція розподілу Ст'юдента для різних n .

Для великих обсягів вибірки розподіл Ст'юдента практично не відрізняється від нормального і практично не залежить від обсягу вибірки. Для малих обсягів вибірки розподіл Ст'юдента дещо відрізняється від нормального, а також залежить від обсягу вибірки.

Якщо $\chi_{n_1}^2$ та $\chi_{n_2}^2$ є незалежними випадковими величинами, що мають розподіл χ^2 з n_1 і n_2 ступенями вільності відповідно, то розподіл випадкової величини

$$F_{n_1, n_2} = \frac{\chi_{n_1}^2 / n_1}{\chi_{n_2}^2 / n_2}$$

називається F – розподілом або розподілом Фішера з n_1 і n_2 ступенями вільності. Густини розподілів χ^2 , Ст'юдента і Фішера можна знайти явно.

Наступні теореми визначають зв'язок між характеристиками вибірки і зазначеними вище розподілами.

Теорема 3. Якщо елементи вибірки $\xi_1, \xi_2, \dots, \xi_n$ незалежні і розподілені нормально з параметрами (a, σ^2) , то вибіркове середнє $\bar{x}$ і вибіркова дисперсія $s^2 = m_2$ незалежні, причому $\bar{x}$ розподілене нормально з параметрами $(a, \sigma^2/n)$, а ns^2/σ^2 має розподіл χ^2 з $n-1$ ступенями вільності.

Теорема 4. Якщо елементи вибірки $\xi_1, \xi_2, \dots, \xi_n$ незалежні і розподілені нормально з параметрами (a, σ^2) , то величина $\frac{\bar{x} - a}{s} \sqrt{n-1}$ має розподіл Ст'юдента з $n-1$ ступенями вільності.

Теорема 5. Якщо $\xi_1, \xi_2, \dots, \xi_{n_1}$ і $\xi_1', \xi_2', \dots, \xi_{n_2}'$ – дві незалежні вибірки і кожна з $n_1 + n_2$ випадкових величин є незалежною і нормально розподіленою з параметрами (a, σ^2) , то випадкова величина $\frac{s^2}{s'^2} \cdot \frac{n_1}{n_2} \cdot \frac{n_2 - 1}{n_1 - 1}$, де $s^2 = m_2$ – вибіркова дисперсія вибірки $\xi_1, \xi_2, \dots, \xi_{n_1}$, а $s'^2 = m_2'$ – вибіркова дисперсія вибірки $\xi_1', \xi_2', \dots, \xi_{n_2}'$, мають розподіл Фішера з $n_1 - 1$ і $n_2 - 1$ степенями свободи. Густина розподілу Фішера має вигляд

$$P_{F_{n_1, n_2}}(x) = \frac{\Gamma\left(\frac{n_1 + n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{x^{\frac{n}{2}-1}}{\left(\frac{n_1}{n_2}x + 1\right)^{\frac{n_1 + n_2}{2}}}, \quad x > 0,$$

$$F_{n_1, n_2} = \frac{\chi_{n_1}^2 / n_1}{\chi_{n_2}^2 / n_2}.$$

Тут всі випадкові величини $\xi_0, \xi_1, \dots, \xi_n$ розподілені за нормальним законом з параметрами $(0,1)$.

3.6. 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 $\zeta < x$

$$F_\zeta^*(x) = \frac{n_x}{n}.$$

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

$$\lim_{n \rightarrow \infty} F_\zeta^*(x) = F_\zeta(x).$$

Для скінчених n

$$F_\zeta^*(x) \approx F_\zeta(x).$$

Емпірична функція розподілу $F_\zeta^*(x)$ має такі властивості:

1. $0 \leq F_\zeta^*(x) \leq 1$;
2. $F_\zeta^*(x)$ – зростаюча функція;
3. $F_\zeta^*(x) = \begin{cases} 0, & x \leq x_{\min}; \\ 1, & x > x_{\max}, \end{cases}$

де $x_{\min}$ – найменше значення елемента вибірки, $x_{\max}$ – найбільше значення елемента вибірки.

Задача 37. Знайти емпіричну функцію розподілу за статистичним розподілом вибірки

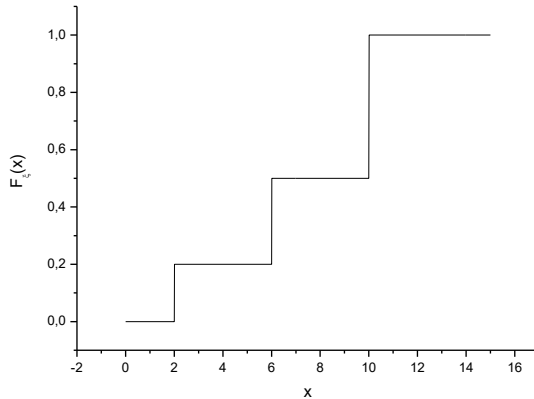
x_i	2	6	10
n_i	12	18	30

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 $x \geq 10$. Значення $\zeta < 6$, тобто $\zeta = \{x_1 = 2\}$, спостерігається 12 разів, тому $F_{\zeta}^*(x) = 12/60 = 0.2$ при $2 \leq x < 6$. 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 $6 \leq x < 10$. Таким чином, одержали емпіричну функцію розподілу вигляду

$$F_{\zeta}^*(x) = \begin{cases} 0, & x < 2; \\ 0.2, & 2 \leq x < 6; \\ 0.5, & 6 \leq x < 10; \\ 1, & x \geq 10. \end{cases}$$

Графік цієї функції наступний



Мал. 20. Емпірична функція розподілу ймовірностей.

Задача 38. У висліді спостереження одержали розподіл ознаки ζ вибірки у вигляді

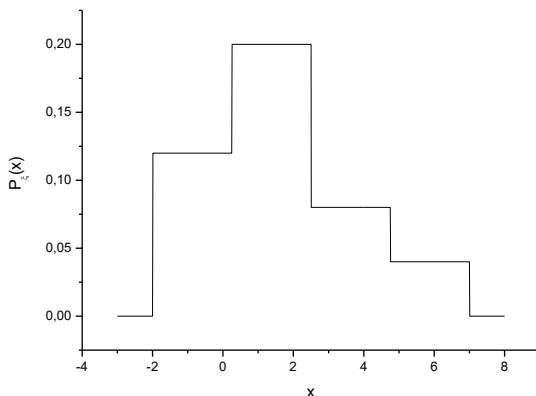
x_k	-2	0	1	2	3	5	7
n_k	4	5	7	8	6	2	1

Побудувати гістограму частот цього розподілу.

Розв'язання. У цьому разі найменше значення випадкової величини $\zeta \in x_{\min} = -2$, 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 будуюмо відповідну гістограму



Мал. 21. Емпірична густина розподілу ймовірностей

Гістограмою частот називають ступінчасту 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 . При цьому площа окремого прямокутника $S_k = \frac{n_k}{nh} h = \frac{n_k}{n}$ якраз дорівнює частоті попадання елемента вибірки у відповідний інтервал. Тут n – обсяг вибірки. Дійсно

$$S = \sum_{k=1}^n S_k = \sum_{k=1}^n \frac{n_k}{nh} h = \frac{1}{n} \sum_{k=1}^n n_k = 1$$

Тобто, як це і випливає з умови нормування густини розподілу, 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 $[x_i, x_{i+1})$, $i = 1, \dots, m$, $x_1 = x_{\min}$, $x_{m+1} = x_{\max}$ 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.

3.7. 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 вибірових характеристик, які також є випадковими величинами. Нижче ми розглянемо саме останню задачу - задачу оцінювання параметрів розподілів.

Нехай випадковий вектор $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_n)$ є вибіркою з генеральної сукупності або вибіровим вектором. Тут $\zeta_1, \zeta_2, \dots, \zeta_n$ – 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 встановлення. 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,

$$\begin{aligned} M\left(\hat{\theta} - \theta\right)^2 &= M\left(\left(\hat{\theta} - M\hat{\theta}\right) + \left(M\hat{\theta} - \theta\right)\right)^2 = \\ &= M\left(\hat{\theta} - M\hat{\theta}\right)^2 + 2M\left(\hat{\theta} - M\hat{\theta}\right)\left(M\hat{\theta} - \theta\right) + M\left(M\hat{\theta} - \theta\right)^2 = \\ &= D\hat{\theta} + 2\left(M\hat{\theta} - M\hat{\theta}\right)\left(M\hat{\theta} - \theta\right) + \left(M\hat{\theta} - \theta\right)^2 = D\hat{\theta} + \left(M\hat{\theta} - \theta\right)^2. \end{aligned}$$

Тут ми врахували, що за означенням дисперсії

$$D\hat{\theta} = M\left(\hat{\theta} - M\hat{\theta}\right)^2$$

$$M\hat{\theta} = \theta, \quad MM^* \hat{\theta} = \hat{\theta}.$$

Отже, якщо $M\hat{\theta} = \theta$, то помилка оцінювання є мінімальною і дисперсію можна вважати її можливою оцінкою. 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.

3.8. Точкова оцінка

Нехай елементи вибірки $x_1, x_2, \dots, x_n$ залежать від невідомого параметра θ . Очевидно від цього параметра буде залежати функція розподілу кожного елемента вибірки

$$P(x_k < x) = F_{\xi}(x, \theta).$$

Оцінкою θ_n^* параметра θ називається довільна функція $\theta_n^* = \theta_n^*(x_1, \dots, x_n)$. 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 $n \rightarrow \infty$

$$P(|\theta_n^* - \theta| < \varepsilon) \rightarrow 1.$$

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

різних оцінок одного і того ж параметра за останнім критерієм використовується наступна величина

$$e(\theta_n^*) = 1 / [nI(\theta)D\theta_n^*],$$

причому

$$0 \leq e(\theta_n^*) \leq 1.$$

Якщо $e(\theta_n^*) = 1$, то оцінку називають ефективною. Тут $D\theta_n^*$ – дисперсія оцінки, а

$$I(\theta) = M \left[\frac{\partial \ln P(x_k, \theta)}{\partial \theta} \right]^2,$$

де $P(x_k, \theta)$ – густина розподілу елемента вибірки x_k .

Задача 39. Нехай $\xi = (x_1, x_2, \dots, x_n)$ – вибірка, що описується нормальним законом з параметрами (a, σ^2) , що є математичним очікуванням і дисперсією окремих елементів вибірки. В якості оцінок a і σ^2 візьмемо вибіркове середнє

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

і виправлену вибірккову дисперсію

$$S^2 = \frac{n}{n-1} m_2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2.$$

Доведемо, що обидві оцінки є незміщеними і слушними.

Розв'язання. Дійсно, оскільки

$$Mx_1 = Mx_2 = \dots = Mx_n = a,$$

то

$$M\bar{x} = M \frac{1}{n}(x_1 + x_2 + \dots + x_n) =$$

$$= \frac{1}{n} (Mx_1 + Mx_2 + \dots + Mx_n) = Mx_i = a.$$

Відповідно, оскільки,

$$Dx_1 = Dx_2 = \dots = Dx_n = \sigma^2,$$

$$m_2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 = \frac{1}{n} \sum_{k=1}^n x_k^2 - \bar{x}^2,$$

$$Mx_k = 0,$$

$$Mx_k x_l = Mx_k Mx_l = 0, \quad (k \neq l),$$

то

$$M\bar{x}^{-2} = M \frac{1}{n^2} \sum_{k,l=1}^n x_k x_l = \frac{1}{n^2} \sum_{k,l=1}^n Mx_k x_l = \frac{1}{n^2} \sum_{k=1}^n Mx_k^2,$$

$$Mm_2 = M \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 = M \left[\frac{1}{n} \sum_{k=1}^n x_k^2 - \bar{x}^2 \right] =$$

$$= \frac{1}{n} \sum_{k=1}^n Mx_k^2 - \frac{1}{n^2} \sum_{k=1}^n Mx_k^2 = \frac{n-1}{n^2} \sum_{k=1}^n Mx_k^2$$

$$MS^2 = \frac{n}{n-1} Mm_2 = \frac{1}{n} \sum_{k=1}^n Mx_k^2 =$$

$$= \frac{1}{n} \sum_{k=1}^n (Mx_k^2 - (Mx_k)^2) = \frac{1}{n} \sum_{k=1}^n Dx_k = \sigma^2.$$

Розглянемо тепер дисперсію цих оцінок

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

Аналогічно можна довести, що

$$DS^2 = \frac{\mu_n - \mu_2^2}{n} + D \left(\frac{1}{n^2} \right).$$

Задача 40. Нехай $\zeta_1, \zeta_2, \dots, \zeta_n$ - вибірка з розподілу з густиною

$$f(x, a, b) = \frac{1}{b-a} \begin{cases} 1, & \text{if } x \in [a, b], \\ 0, & \text{if } x \notin [a, b]. \end{cases}$$

Які з наступних оцінок та яких параметрів є незміщеними оцінками?

$$\hat{\theta}_1 = \min \{ \zeta_1, \zeta_2, \dots, \zeta_n \};$$

$$\hat{\theta}_2 = \max \{ \zeta_1, \zeta_2, \dots, \zeta_n \};$$

$$\hat{\theta}_3 = \frac{1}{n} \sum_{i=1}^n \zeta_i;$$

$$\hat{\theta}_n = (\zeta_{n-1} + \zeta_n) / 2.$$

Розв'язання. Спочатку знайдемо розподіл випадкової величини $\hat{\theta}_2 = \max \{ \zeta_1, \zeta_2, \dots, \zeta_n \}$. Оскільки всі випадкові величини $\zeta_1, \zeta_2, \dots, \zeta_n$ незалежні, то, відповідно до теореми множення ймовірностей

$$\begin{aligned} F_{\hat{\theta}_2}(x) &= P(\hat{\theta}_2 < x) = P\{\max \{ \zeta_1, \zeta_2, \dots, \zeta_n \} < x\} = \\ &= P\{\zeta_1 < x \cap \zeta_2 < x \cap \dots \cap \zeta_n < x\} = \prod_{i=1}^n P\{\zeta_i < x\}. \end{aligned}$$

Тепер врахуємо, що всі випадкові величини $\zeta_1, \zeta_2, \dots, \zeta_n$ однаково розподілені, тобто

$$P\{\zeta_1 < x\} = P\{\zeta_2 < x\} = \dots = P\{\zeta_n < x\}$$

або

$$F_{\zeta_1}(x) = F_{\zeta_2}(x) = \dots = F_{\zeta_n}(x).$$

Тоді

$$F_{\hat{\theta}_2}(x) = \prod_{i=1}^n F_{\zeta_i}(x) = [F_{\zeta}(x)]^n = \left[\int_{-\infty}^x f(y, a, b) dy \right]^n.$$

За відомою функцією розподілу випадкової величини $\hat{\theta}_2$ знайдемо густину її розподілу. За означенням

$$f_{\hat{\theta}_2}(x) = \frac{d}{dx} F_{\hat{\theta}_2}(x) = n f(x, a, b) \left[\int_{-\infty}^x f(y, a, b) dy \right]^{n-1}.$$

Оскільки

$$\int_{-\infty}^x f(y, a, b) dy = \begin{cases} 0, & \text{if } x < a, \\ \frac{x-a}{b-a}, & \text{if } a \leq x \leq b, \\ 1, & \text{if } x > b. \end{cases}$$

Отже,

$$f_{\hat{\theta}_2}(x) = \begin{cases} \frac{n}{(b-a)^n} (x-a)^{n-1}, & \text{if } x \in [a, b], \\ 0, & \text{if } x \notin [a, b]. \end{cases}$$

За відомою щільністю розподілу оцінки $\hat{\theta}_2$ можемо обчислити її математичне очікування:

$$M\hat{\theta}_2 = \int_{-\infty}^{\infty} x f_{\hat{\theta}_2}(x) dx = \int_a^b x \frac{n}{(b-a)^n} (x-a)^{n-1} dx = \frac{n}{n+1} b + \frac{a}{n+1}.$$

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

$$\lim_{n \rightarrow \infty} M \hat{\theta}_2 = \lim_{n \rightarrow \infty} \left[\frac{n}{n+1} b + \frac{a}{n+1} \right] = b.$$

Знайдемо тепер розподіл випадкової величини $\hat{\theta}_1 = \min \{ \zeta_1, \zeta_2, \dots, \zeta_n \}$, тобто функцію

$$F_{\hat{\theta}_1}(x) = P \{ \hat{\theta}_1 < x \} = P \{ \min \{ \zeta_1, \zeta_2, \dots, \zeta_n \} < x \}.$$

Якщо за аналогією з попереднім розглядом перейти далі від розподілу ймовірностей для випадкової величини $\hat{\theta}_1$ до функцій розподілу випадкових величин $\zeta_1, \zeta_2, \dots, \zeta_n$, то виникає незручність, пов'язана з тим, що з умови $\min \{ \zeta_1, \zeta_2, \dots, \zeta_n \} < x$ не випливають умови $\zeta_1, \zeta_2, \dots, \zeta_n < x$. Оскільки

$$P \{ \hat{\theta}_1 < x \} + P \{ \hat{\theta}_1 \geq x \} = 1,$$

то

$$F_{\hat{\theta}_1}(x) = 1 - P \{ \hat{\theta}_1 \geq x \} = 1 - \prod_{i=1}^n P \{ \zeta_i \geq x \} = 1 - \left[\int_x^{\infty} f(y, a, b) dy \right]^n.$$

Обчислимо інтеграл

$$\int_x^{\infty} f(y, a, b) dy = \begin{cases} 1, & \text{if } x < a, \\ \frac{b-x}{b-a}, & \text{if } a \leq x \leq b, \\ 0, & \text{if } x > b. \end{cases}$$

Тепер

$$f_{\hat{\theta}_1}(x) = \frac{d}{dx} F_{\hat{\theta}_1}(x) = n f(x, a, b) \left[\int_x^{\infty} f(y, a, b) dy \right]^{n-1} =$$

$$= \begin{cases} \frac{n}{(b-a)^n}, \text{ if } x \in [a, b], \\ 0, \text{ if } x \notin [a, b]. \end{cases}$$

За відомою щільністю розподілу оцінки $\hat{\theta}_1$ обчислимо її математичне очікування

$$M\hat{\theta}_1 = \int_{-\infty}^{\infty} x f_{\hat{\theta}_1}(x) dx = \int_a^b x \frac{n}{(b-a)^n} (b-x)^{n-1} dx = \frac{n}{n+1} a + \frac{b}{n+1},$$

тобто $\hat{\theta}_1$ не є незміщеною оцінкою ні параметра a , ні параметра b , але вона є асимптотично незміщеною оцінкою параметра a . Дійсно

$$\lim_{n \rightarrow \infty} M\hat{\theta}_1 = \lim_{n \rightarrow \infty} \left[\frac{n}{n+1} a + \frac{b}{n+1} \right] = a.$$

Обчислимо математичне очікування оцінки $\hat{\theta}_3$. Оскільки випадкова величина $\hat{\theta}_3$ є сумою випадкових величин $\zeta_1, \zeta_2, \dots, \zeta_n$, 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 очікування 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 очікування рівномірного на відрізку $[a, b]$ розподілу. Аналогічно попередньому

$$M\hat{\theta}_n = \frac{M(\zeta_{n-1} + \zeta_n)}{2} = \frac{a+b}{2},$$

тобто оцінка $\hat{\theta}_n$ також є незміщеною оцінкою математичного очікування 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 очікування 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

$$\begin{aligned} P\left\{\left|\hat{\theta}_2 - b\right| > \varepsilon\right\} &= P\left\{\hat{\theta}_2 \in (-\infty, b - \varepsilon) \cup (b + \varepsilon, \infty)\right\} = \\ &= \int_{-\infty}^{b-\varepsilon} f_{\hat{\theta}_2}(x) dx + \int_{b+\varepsilon}^{\infty} f_{\hat{\theta}_2}(x) dx = \int_a^{b-\varepsilon} f_{\hat{\theta}_2}(x) dx = \\ &= \int_a^{b-\varepsilon} \frac{n}{(b-a)^n} (x-a)^{n-1} dx = \left(1 - \frac{\varepsilon}{b-a}\right)^n. \end{aligned}$$

Звідси

$$\lim_{n \rightarrow \infty} P \left\{ \left| \hat{\theta}_2 - b \right| > \varepsilon \right\} = \lim_{n \rightarrow \infty} \left(1 - \frac{\varepsilon}{b-a} \right)^n = 0$$

і оцінка $\hat{\theta}_2$ є слушною оцінкою параметра θ . Так само слушною є і оцінка θ . Слушність оцінки $\hat{\theta}_3$ впливає 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 неслухною.

3.8.1. 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 $x_1, x_2, \dots, x_n$, де кожному елементу вибірки відповідає одна і та сама густина ймовірності $P_\xi(x, \theta)$, то густиною ймовірності для випадкового вектора $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ буде функція $P_\xi(x, \theta) = P_{\xi_1}(x, \theta) = \dots = P_{\xi_n}(x, \theta)$. Максимальне значення ця густина ймовірності матиме для $\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n$. Таке максимальне значення називається функцією правдоподібності

$$L = L(x_1, x_2, \dots, x_n, \theta) = P_{\xi_1}(x_1, \theta) P_{\xi_2}(x_2, \theta) \cdots P_{\xi_n}(x_n, \theta).$$

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

$$\frac{\partial \ln(L)}{\partial \theta} = 0.$$

Розв'язки цього рівняння називаються оцінками найбільшої правдоподібності для параметру θ .

Якщо невідомих параметрів декілька. Нехай це $\theta_1, \theta_2, \dots, \theta_r$, то замість одного рівняння правдоподібності, слід розв'язати систему таких рівнянь

$$\frac{\partial \ln(L)}{\partial \theta_k} = 0, k = 1, 2, \dots, r.$$

Приклад 1. Нехай випадкові величини $\xi_1, \xi_2, \dots, \xi_k$, які ми співставили елементам вибірки $x_1, x_2, \dots, x_k$ мають нормальний розподіл, і невідомими параметрами є їх математичне очікування $a = M\xi_k$ та дисперсія $b = \sigma^2 = D\xi_k$. Знайти їх оцінку найбільшої правдоподібності.

Розв'язання. Функція найбільшої правдоподібності має вигляд

$$L = L(x_1, \dots, x_n, a, b) = \left(\frac{1}{\sqrt{2\pi b}} \right)^n \exp \left(- \sum_{k=1}^n \frac{(x_k - a)^2}{2b} \right).$$

Відповідно

$$\ln(L) = -\frac{n}{2}(\ln 2\pi + \ln b) - \frac{1}{2b} \sum_{k=1}^n (x_k - a)^2.$$

Для оцінок a і b параметрів a, b отримаємо наступну систему рівнянь правдоподібності

$$\frac{\partial \ln(L)}{\partial a} = \frac{1}{\tilde{b}} \sum_{k=1}^n (x_k - a) = 0,$$

$$\frac{\partial \ln(L)}{\partial b} = -\frac{n}{2\tilde{b}} + \frac{1}{2\tilde{b}^2} \sum_{k=1}^n (x_k - a)^2 = 0.$$

вибірки. Перше рівняння можна записати так

$$\sum_{k=1}^n x_k - na = 0.$$

Звідси

$$a = \bar{a} = \frac{1}{n} \sum_{k=1}^n x_k.$$

Для другого рівняння маємо

$$-n + \frac{1}{\tilde{b}} \sum_{k=1}^n (x_k - a)^2 = 0$$

Звідки

$$\tilde{b} = \bar{b} = s^2 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2.$$

Отже, вибіркове середнє є найкращою в сенсі методу найбільшої правдоподібності оцінкою математичного очікування, а вибіркова дисперсія є аналогічною оцінкою дисперсії.

Приклад 2. Нехай випадкові величини $\xi_1, \xi_2, \dots, \xi_k$, які ми співставили елементам вибірки $x_1, x_2, \dots, x_k$, мають закон розподілу Бернуллі, а невідомим параметром є ймовірність успіху p . Знайти його оцінку методом 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

$$L = L(m_1, \dots, m_n, p) = \left(C_n^m\right)^n p^{m_1 + \dots + m_n} (1-p)^{n^2 - (m_1 + \dots + m_n)}.$$

Відповідно

$$\ln(L) = L = n \ln C_n^m + (m_1 + \dots + m_n) \ln p + (n^2 - (m_1 + \dots + m_n)) \ln(1-p).$$

Для оцінки p параметру p отримаємо наступне рівняння правдоподібності вибірки

$$\frac{\partial \ln(L)}{\partial p} = \frac{m_1 + \dots + m_n}{p} - \frac{n^2 - (m_1 + \dots + m_n)}{1-p} = 0$$

або

$$\frac{\bar{m}}{p} - \frac{n - \bar{m}}{1-p} = 0$$

Це рівняння має очевидний розв'язок

$$p = \frac{\bar{m}}{n},$$

де

$$\bar{m} = \frac{m_1 + \dots + m_n}{n}$$

- вибіркове середнє кількості успіхів.

3.8.2. 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:

$$a_v = \frac{1}{n} \sum_{i=1}^n x_i^v$$

- вибіркові початкові моменти,

$$m_v = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^v$$

- центральні вибіркові моменти,

$$\alpha_v(\lambda) = M \xi^v = \int x^v dF_{\xi}(x, \lambda)$$

- теоретичні початкові моменти для неперервної випадкової величини та

$$\mu_{\alpha}(\lambda) = M (\xi - M \xi)^v = \int (x - M \xi)^v dF_{\xi}(x, \lambda)$$

- 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: $\lambda_1, \lambda_2, \dots, \lambda_k$, то слід прирівняти відповідну кількість перших моментів

$$a_1 = \alpha_1(\lambda_1, \lambda_2, \dots, \lambda_k),$$

$$a_2 = \alpha_2(\lambda_1, \lambda_2, \dots, \lambda_k),$$

.....

$$a_k = \alpha_k(\lambda_1, \lambda_2, \dots, \lambda_k)$$

або

$$m_1 = \mu_1(\lambda_1, \lambda_2, \dots, \lambda_k),$$

$$m_2 = \mu_2(\lambda_1, \lambda_2, \dots, \lambda_k),$$

.....

$$m_k = \mu_k(\lambda_1, \lambda_2, \dots, \lambda_k)$$

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

$$a_1 = \alpha_1(a, \sigma),$$

$$m_2 = \mu_2(a, \sigma).$$

Оскільки

$$a_1 = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$m_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

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

$$a = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$\sigma^2 = s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

3.9. 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 $x_1, x_2, \dots, x_n$ з функцією розподілу $P(x_i < x) = F_\xi(x, \theta)$, де θ – невідомий параметр. Інтервал $(\underline{\theta}, \bar{\theta})$, у якому із заданою імовірністю 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.

3.9.1. Довірчий інтервал для математичного очікування нормального розподілу при відомій дисперсії

Нехай елементи вибірки розподілені нормально з параметрами (a, σ^2) . При цьому параметр a невідомий, а параметр σ^2 відомий. Знайдемо довірчий інтервал для a . Випадкова величина (вибіркове середнє) $\bar{x} = (x_1 + \dots + x_n) / n$, як відомо, у зазначеному випадку теж має нормальний розподіл з параметрами $(a, \sigma^2 / n)$, тобто $M\bar{x} = a$, $D\bar{x} = \sigma^2 / n$. Оскільки функція розподілу ймовірностей для вибірки залежить від невідомого параметру a , то побудувати довірчий інтервал для цього параметру, використовуючи таку функцію, незручно. У разі нормального розподілу легко сконструювати випадкову величину,

функція розподілу ймовірностей якої не залежить від шуканого параметру. Ця залежність переноситься на саму випадкову величину. Такою випадковою величиною є величина $(\bar{x} - a)/(\sigma/\sqrt{n})$, розподілена нормально з параметрами $(0,1)$.

Очевидно, її розподіл вже не залежить від a . Звідси

$$P\left\{\left|\frac{\bar{x} - a}{\sigma/\sqrt{n}}\right| < u_{\alpha}\right\} = 1 - 2\alpha,$$

або

$$P\left\{\bar{x} - u_{\alpha} \frac{\sigma}{\sqrt{n}} < a < \bar{x} + u_{\alpha} \frac{\sigma}{\sqrt{n}}\right\} = 1 - 2\alpha,$$

де $1 - 2\alpha$ - довірча ймовірність, $\left[\bar{x} - u_{\alpha} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{\alpha} \frac{\sigma}{\sqrt{n}}\right]$ - довірчий

інтервал, тобто інтервал, у якому з довірчою ймовірністю лежить можливе значення шуканого параметру. Довірча ймовірність звичайно мало відрізняється від одиниці і її запис у вигляді $1 - 2\alpha$ має на увазі, що 2α є малою величиною. Двійка у цьому виразі звичайно пишеться для того, щоб величину α зручніше виразити через значення спеціальної функції - функції Лапласа. Зазначимо, що u_{α} є величиною відхилення вибіркового середнього від значення оцінюваного параметра a , виражене в одиницях середньоквадратичного відхилення для вибіркового середнього $\sigma/\sqrt{n}$.

Ще раз наголошуємо, що у разі нормального розподілу ні межі довірчого інтервалу, ні довірча ймовірність не залежать від шуканого параметру. Для випадку нормального розподілу дуже легко записати зв'язок між α і u_{α} . Дійсно, для нормального розподілу ймовірностей

$$P\left\{\left|\frac{\bar{x}-a}{\sigma/\sqrt{n}}\right| < u_{\alpha}\right\} = 1 - 2\alpha = \frac{1}{\sqrt{2\pi}} \int_{-u_{\alpha}}^{u_{\alpha}} \exp(-y^2/2) dy =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-u_{\alpha}}^0 \exp(-y^2/2) dy + \frac{1}{\sqrt{2\pi}} \int_0^{u_{\alpha}} \exp(-y^2/2) dy = 2\Phi_0(u_{\alpha}),$$

де

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

- функція Лапласа ($\Phi_0(\infty) = 1/2$). Тут ми використали ту обставину, що

$$\int_{-u_{\alpha}}^0 \exp(-y^2/2) dy = - \int_0^{-u_{\alpha}} \exp(-y^2/2) dy = \int_0^{u_{\alpha}} \exp(-y^2/2) dy.$$

Отже,

$$P\left\{\bar{x} - u_{\alpha} \frac{\sigma}{\sqrt{n}} < a < \bar{x} + u_{\alpha} \frac{\sigma}{\sqrt{n}}\right\} = 2\Phi_0(u_{\alpha})$$

або

$$\alpha = 1/2 - \Phi_0(u_{\alpha}).$$

Останній результат часто наводять і у наступному вигляді

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_{u_{\alpha}}^{\infty} \exp(-y^2/2) dy.$$

Для практичного використання цих формул звичайно використовують як таблиці значень функції Лапласа $\Phi_0(u_{\alpha})$, так і таблиці функції $1/2 - \Phi_0(u_{\alpha})$.

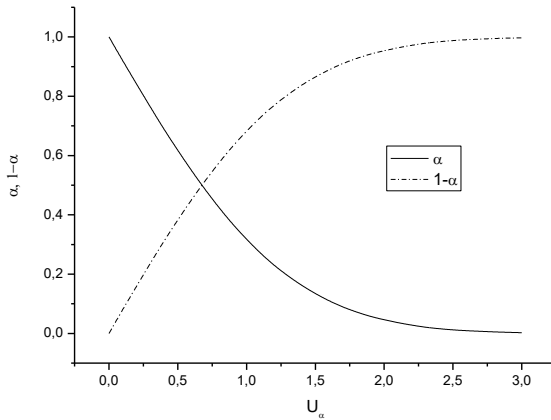
Варіантів реалізації практичних формул для даного випадку досить багато. Зустрічається і такий варіант

$$P\left\{\left|\frac{\bar{x}-a}{\sigma/\sqrt{n}}\right| < u_{\alpha}\right\} = 1 - \alpha = 2\Phi_0(u_{\alpha}).$$

Звідки

$$1 - 2\Phi_0(u_\alpha) = \alpha$$

Графічно цей зв'язок виглядає наступним чином.



Мал. 22. Залежність довірчої ймовірності від довірчого інтервалу у разі нормального розподілу для математичного очікування.

Якщо $u_\alpha = 1$, то

$$P\left\{\bar{x} - \frac{\sigma}{\sqrt{n}} < a < \bar{x} + \frac{\sigma}{\sqrt{n}}\right\} = 0.6826.$$

Якщо $u_\alpha = 2$, то

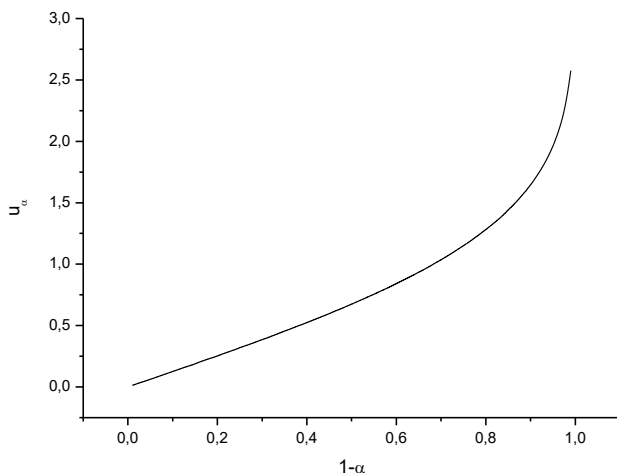
$$P\left\{\bar{x} - 2\frac{\sigma}{\sqrt{n}} < a < \bar{x} + 2\frac{\sigma}{\sqrt{n}}\right\} = 0.9548.$$

Якщо $u_\alpha = 3$, то

$$P\left\{\bar{x} - 3\frac{\sigma}{\sqrt{n}} < a < \bar{x} + 3\frac{\sigma}{\sqrt{n}}\right\} = 0.9972.$$

Майже достовірним є той факт, що випадкова величина не відхилиться від математичного очікування (вибіркового середнього) більше, ніж на 3σ .

Конкретну залежність ширини довірчого інтервалу від довірчої ймовірності для нормального розподілу з нульовим математичним очікуванням $\mu = 0$ і одиничною дисперсією $\sigma = 1$ можна побачити на наступному графіку



Мал. 23. Залежність розмірів довірчого інтервалу від довірчої ймовірності для математичного очікування.

З останнього малюнка видно, що ширина довірчого інтервалу швидко зростає із зростанням довірчої ймовірності але при ширині $3\sigma = 3$ довірна ймовірність вже практично досягає свого максимального значення, рівного 1.

Задача 41. Нехай випадкова величина набуває наступних значень: 3.1, 3.3, 2.9, 3.0, 3.1, 3.2, 2.8, 2.7, 3.1, 3.2, 2.9, 3.0, 2.9, 3.1, 2.8, 2.9, 3.2, 3.3, 2.9, 3.1, 3.2, 3.0. Дисперсія цієї випадкової величини

дорівнює $\sigma^2 = 0.028$. Визначити залежність довірчого інтервалу від довірчої ймовірності для математичного очікування.

Розв'язання. Знайдемо вибіркове середнє

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Для $n=22$ $\bar{x}=3.032$. Для меж довірчого інтервалу $[a_\alpha, b_\alpha]$ маємо:

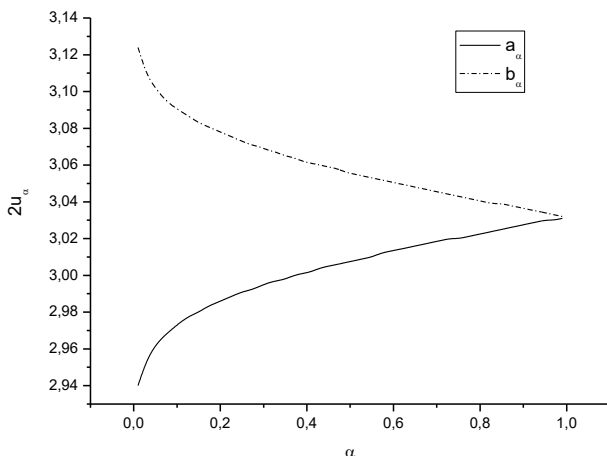
$$a_\alpha = \bar{x} - \frac{\sigma}{\sqrt{n}} u_\alpha,$$

$$b_\alpha = \bar{x} + \frac{\sigma}{\sqrt{n}} u_\alpha.$$

Тут межі довірчого інтервалу для стандартної нормальної функції розподілу з параметрами (0,1) визначаються з рівняння, що виражає довірчу ймовірність через довірчий інтервал

$$1 - 2\alpha = 2\Phi_0(u_\alpha).$$

Розв'язок цього рівняння для різних значень довірчої ймовірності можна знайти або за таблицями, що часто присутні в підручниках з математичної статистики, або використанням математичних прикладних пакетів типу Mathcad, Matlab тощо. У висліді отримуємо таку графічну залежність



Мал. 24. Залежність ширини довірчого інтервалу від довірчої ймовірності для математичного очікування

З графіку видно, що ширина довірчого інтервалу швидко зменшується при зростанні α або при зменшенні довірчої ймовірності $1 - 2\alpha$. Нагадаємо, що 2α - це ймовірність того, що шукана величина лежить поза межами довірчого інтервалу.

3.9.2. Довірчий інтервал для математичного очікування нормального розподілу 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 $\bar{x}$ і S^2 (вибіркового середнього і виправленої вибіркової дисперсії) можна утворити третю випадкову величину

$$\tau_{n-1} = \frac{\bar{x} - a}{S / \sqrt{n-1}}.$$

Як відомо, для елементів вибірки розподілених за нормальним законом, випадкова величина τ_{n-1} розподілена за законом Ст'юдента з $n-1$ степенями свободи. Також, як і у разі нормального розподілу, для введеної випадкової величини, що описується розподілом Ст'юдента, ні межі довірчого інтервалу, ні довірча ймовірність не залежать від шуканих параметрів. Натомість з'являється залежність від обсягу вибірки n . Для випадку розподілу Ст'юдента також дуже легко записати зв'язок між довірчим інтервалом і довірчою ймовірністю, зокрема між α і $t_{\alpha, n-1}$. Дійсно, нехай

$$P\{|\tau_{n-1}| < t_{\alpha, n-1}\} = 1 - \alpha,$$

або

$$P\left\{\bar{x} - t_{\alpha, n-1} \frac{S}{\sqrt{n-1}} < a < \bar{x} + t_{\alpha, n-1} \frac{S}{\sqrt{n-1}}\right\} = 1 - \alpha.$$

Зв'язок між величинами $t_{\alpha, n-1}$ і α визначається законом розподілу Ст'юдента

$$P\{|\tau_{n-1}| < t_{\alpha, n-1}\} = 1 - \alpha = \int_{-t_{\alpha, n-1}}^{t_{\alpha, n-1}} P_{\tau_{n-1}}(x) dx.$$

Відповідно,

$$P\{|\tau_{n-1}| > t_{\alpha, n-1}\} = \alpha = 1 - \int_{-t_{\alpha, n-1}}^{t_{\alpha, n-1}} P_{\tau_{n-1}}(x) dx$$

або

$$\alpha = 2 \int_{t_{\alpha, n-1}}^{\infty} P_{\tau_{n-1}}(x) dx = 2 \int_{-\infty}^{-t_{\alpha, n-1}} P_{\tau_{n-1}}(x) dx.$$

Наприклад, для довірчої ймовірності $1-\alpha=0.99$, $\alpha=0.01$ і $n-1=10$ отримаємо $t_{\alpha,n-1}=3.169$.

Розподіл Ст'юдента переходить в нормальний розподіл у межі $n \rightarrow \infty$. На практиці це означає, що для вибірок обсягом більше тридцяти можна використовувати нормальний розподіл ймовірності, для менших вибірок - розподіл Ст'юдента.

Густина розподілу Ст'юдента з n степенями вільності має вигляд

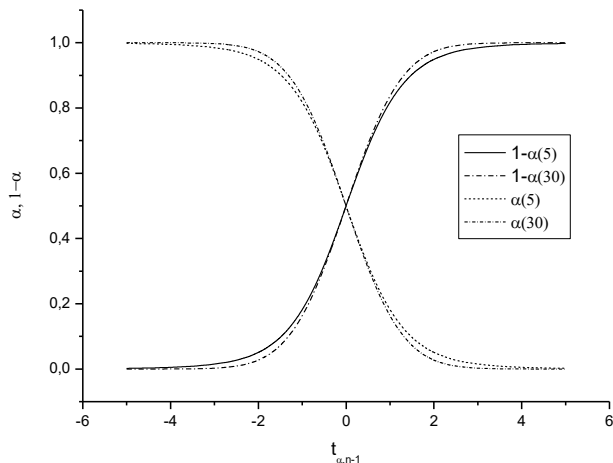
$$P_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in (-\infty, \infty).$$

Отже, враховуючи, що маємо $n-1$ ступінь вільності,

$$\alpha = 2 \int_{t_{\alpha,n-1}}^{\infty} \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)} \Gamma\left(\frac{n-1}{2}\right)} \left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}} dx,$$

$$1-\alpha = \int_{-\infty}^{t_{\alpha,n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)} \Gamma\left(\frac{n-1}{2}\right)} \left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}} dx.$$

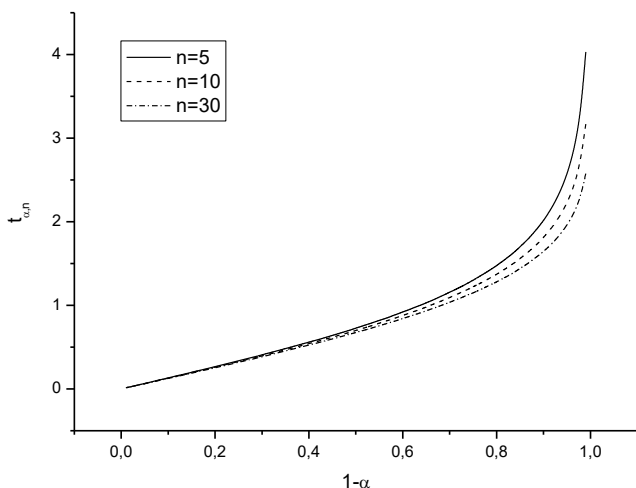
Графічно залежність між довірчим ймовірністю і довірчим інтервалом можна побачити на наступному малюнку



Мал. 25. Зв'язок довірчої ймовірності і довірчого інтервалу у разі розподілу Ст'юдента для математичного очікування.

З малюнка видно, що різниця між довірчим інтервалом і довірчою ймовірністю у разі розподілу Ст'юдента не надто відрізняється від випадку нормального розподілу. Для кількості ступенів вільності $n = 30$ їх можна вважати однаковими і використовувати будь-який з них. У разі $n = 5$ слід використовувати саме розподіл Ст'юдента.

Можна також відобразити і обернену залежність - довірчого інтервалу від довірчої ймовірності.



Мал. 26. *Зв'язок довірчого інтервалу і довірчої ймовірності у разі розподілу Ст'юдента для математичного очікування.*

Випадок 30 ступенів вільності графічно відрізнити від випадку нормального розподілу практично неможливо.

Всі наведені оцінки правильні 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 Ст'юдента 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 гранична теорема. Дійсно, відповідно до неї, незалежно від функції розподілу елементів вибірки, вибіркове середнє при досить великому обсягу вибірки описуватиметься нормальним законом розподілу, а вибіркова дисперсія - законом Ст'юдента.

Задача 42. Нехай випадкова величина набуває наступних значень: 3.1, 3.3, 2.9, 3.0, 3.1, 3.2, 2.8, 2.7, 3.1, 3.2, 2.9, 3.0, 2.9, 3.1, 2.8, 2.9, 3.2, 3.3, 2.9, 3.1, 3.2, 3.0. Дисперсія цієї випадкової величини невідома. Визначити залежність довірчого інтервалу від довірчої ймовірності для математичного очікування.

Розв'язання. Знайдемо вибіркове середнє і виправлену вибірккову дисперсію

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Для $n = 22$ $\bar{x} = 3.032$, $S^2 = 0.028$. Для меж довірчого інтервалу $[a_{\alpha,n}, b_{\alpha,n}]$ маємо:

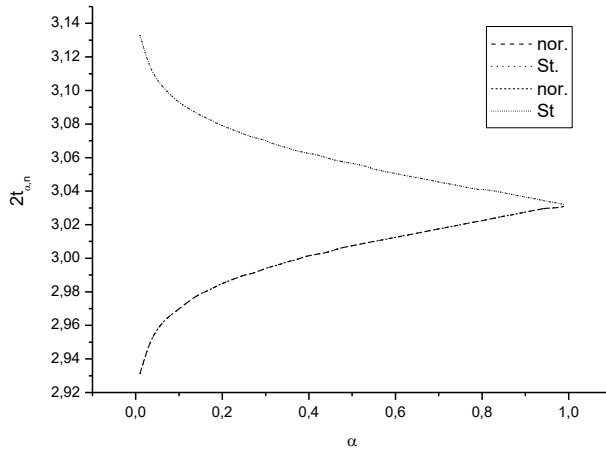
$$a_{\alpha,n-1} = \bar{x} - t_{\alpha,n-1} \frac{S}{\sqrt{n-1}},$$

$$b_{\alpha,n-1} = \bar{x} + t_{\alpha,n-1} \frac{S}{\sqrt{n-1}}.$$

Тут межі довірчого інтервалу функції розподілу Ст'юдента з 21 ступенем вільності визначаються з рівняння, що пов'язує довірчу ймовірність і довірчий інтервал

$$1 - \alpha = \int_{-t_{\alpha,n-1}}^{t_{\alpha,n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)} \Gamma\left(\frac{n-1}{2}\right)} \left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}} dx.$$

Розв'язок цього рівняння для різних значень довірчої ймовірності можна знайти або за таблицями, що часто присутні в підручниках з математичної статистики, або використанням математичних прикладних пакетів типу Mathcad, Matlab тощо. У висліді отримуємо таку графічну залежність



Мал. 27. Залежність ширини довірчого інтервалу від довірчої ймовірності для математичного очікування.

З графіку видно, що ширина довірчого інтервалу швидко зменшується при зростанні α або при зменшенні довірчої ймовірності $1-\alpha$. Нагадаємо, що α - це ймовірність того, що шукана величина лежить поза межами довірчого інтервалу. Крім того, для вибірки обсягом 21 на малюнку практично неможливо розрізнити графіки нормального розподілу і розподілу Ст'юдента.

3.9.3. Довірчий інтервал для дисперсії нормального розподілу при відомому математичному очікуванні

Розглянемо тепер ситуацію, коли для вибірки $x_1, x_2, \dots, x_n$ 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

$$\chi_{n-1}^2 = (n-1)S^2 / \sigma^2$$

розподілена за законом χ^2 з $n-1$ ступенем вільності. Тут виправлена вибіркова дисперсія

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Для розподілу χ^2 також дуже легко записати зв'язок між довірчим інтервалом і довірчою ймовірністю. Дійсно, нехай

$$P\{\chi_1^2 < (n-1)S^2 / \sigma^2 < \chi_2^2\} = 1 - \alpha,$$

або

$$P\{(n-1)S^2 / \chi_2^2 < \sigma^2 < (n-1)S^2 / \chi_1^2\} = 1 - \alpha$$

або

$$P\{\sqrt{n-1}S / \chi_2 < \sigma < \sqrt{n-1}S / \chi_1\} = 1 - \alpha.$$

Таким чином, інтервал $((n-1)S^2 / \chi_2^2, (n-1)S^2 / \chi_1^2)$ є довірчим інтервалом для дисперсії, а інтервал $(\sqrt{n-1}S / \chi_2, \sqrt{n-1}S / \chi_1)$ є довірчим інтервалом для середньоквадратичного відхилення. Використовуючи густину розподілу χ^2

$$P_{\chi_n^2}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad x > 0, \quad \chi_n^2 = \xi_1^2 + \dots + \xi_n^2$$

можна записати наступний зв'язок між довірчим інтервалом і довірчою ймовірністю

$$1 - \alpha = \int_{\chi_1^2}^{\chi_2^2} P_{\chi_{n-1}^2}(x) dx$$

або

$$1 - \alpha = \int_{\chi_1^2}^{\chi_2^2} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx .$$

Очевидно, що числа χ_1^2 і χ_2^2 можна обрати нескінченною кількістю способів. Певну визначеність тут можуть внести міркування симетрії. Розглянемо їх. Числа χ_1^2 і χ_2^2 розбивають проміжок $[0, \infty)$ на три проміжки: $[0, \chi_1^2)$, $[\chi_1^2, \chi_2^2]$ і (χ_2^2, ∞) . Приналежності дисперсії довірчому інтервалу $[\chi_1^2, \chi_2^2]$ відповідає довірча ймовірність $1 - \alpha$. Двом іншим інтервалам відповідає ймовірність α . Доцільно цю останню ймовірність поділити порівну між цими двома інтервалами, тобто покласти

$$\frac{\alpha}{2} = \int_0^{\chi_1^2} P_{\chi_{n-1}^2}(x) dx ,$$

$$\frac{\alpha}{2} = \int_{\chi_2^2}^{\infty} P_{\chi_{n-1}^2}(x) dx .$$

Задача 43. Нехай випадкова величина набуває наступних значень: 3.1, 3.3, 2.9, 3.0, 3.1, 3.2, 2.8, 2.7, 3.1, 3.2, 2.9, 3.0, 2.9, 3.1, 2.8, 2.9, 3.2, 3.3, 2.9, 3.1, 3.2, 3.0. Дисперсія цієї випадкової величини невідома. Визначити залежність довірчого інтервалу від довірчої ймовірності для дисперсії.

Розв'язання. Знайдемо вибіркове середнє і виправлену вибірккову дисперсію

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i ,$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 .$$

Для $n = 22$ $\bar{x} = 3.032$, $S^2 = 0.028$. Для меж довірчого інтервалу $[a_{\alpha,n}, b_{\alpha,n}]$ маємо

$$a_{\alpha n-1} = (n-1)S^2 / \chi^2_2 ,$$

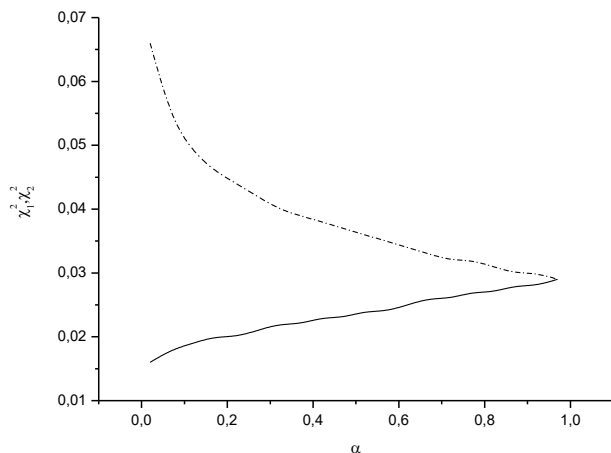
$$b_{\alpha n-1} = (n-1)S^2 / \chi^2_1 .$$

Тут межі довірчого інтервалу функції розподілу χ^2 з $n-1$ ступенем вільності визначаються з рівнянь, що пов'язують довірчу ймовірність з межами довірчого інтервалу

$$\frac{\alpha}{2} = \int_0^{\chi^2_1} P_{\chi^2_{n-1}}(x) dx ,$$

$$\frac{\alpha}{2} = \int_{\chi^2_2}^{\infty} P_{\chi^2_{n-1}}(x) dx ..$$

Розв'язок цього рівняння для різних значень довірчої ймовірності можна знайти або за таблицями, що часто присутні в підручниках з математичної статистики, або використанням математичних прикладних пакетів типу Mathcad, Matlab тощо. У висліді отримуємо таку графічну залежність



Мал. 28. Залежність ширини довірчого інтервалу від довірчої ймовірності у разі розподілу χ^2 для дисперсії.

З графіку видно, що ширина довірчого інтервалу швидко зменшується при зростанні α або при зменшенні довірчої ймовірності $1-\alpha$. Нагадаємо, що α - це ймовірність того, що шукана величина лежить поза межами довірчого інтервалу.

3.10. Статистична перевірка гіпотез. Критерій Пірсона

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 $S_1, S_2, \dots, S_r$ без спільних точок. Для цього нам потрібно $r-1$ число $a_1, a_2, \dots, a_{r-1}$. При цьому правий кінець відповідного інтервалу виключається з відповідної множини, а лівий включається. Нехай $p_i, i = 0, 1, \dots, r$ теоретичні ймовірності того, що випадкова величина ξ належить

множині S_i . Очевидно, що $\sum_{i=1}^r p_i = 1$. Нехай 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 величина

$$\chi^2 = \sum_{i=1}^r \frac{n}{p_i} \left(\frac{n_i}{n} - p_i \right)^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}$$

може служити критерієм близькості теоретичної і експериментальної функцій розподілу. Тут у чисельнику

міститься квадрат відхилення кількості елементів вибірки n_i , що потрапили в множину S_i і теоретичної кількості елементів, що мали б туди потрапити np_i , а в знаменнику - теоретична кількість елементів np_i . Тобто мова йде про відносну величину квадрату відхилення. Привабливою рисою такого критерію є те, що структура правої частини останньої рівності відповідає структурі випадкової величини, розподіленої за законом χ^2 , що стає цілком зрозумілим з теореми, поданої нижче

$$\chi^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_r^2,$$

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 частоти 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 χ^2 ? Як функція випадкових величин, вона так само є випадковою величиною. Можливість порівняння виникла б лише у разі, якщо функція розподілу випадкової величини χ^2 була б відома.

Пірсон не лише запропонував критеріальний вираз але і довів теорему про те, що за достатньо великого обсягу вибірки

випадкова величина χ^2 описується розподілом ймовірності χ^2 з $r - 1$ ступенем вільності, тобто

$$P(\chi^2 < \chi_0^2) = \int_0^{\chi_0^2} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx,$$

$$P(\chi^2 > \chi_0^2) = \int_{\chi_0^2}^{\infty} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx.$$

Задаючи ймовірність $P(\chi^2 < \chi_0^2)$, ми отримуємо теоретичну можливість знаходження граничного значення χ_0^2 . Використовуючи критерій Пірсона, ми отримуємо експериментальну можливість обчислення величини χ^2 . Якщо $\chi^2 < \chi_0^2$, то значення величини χ^2 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 $\chi^2 > \chi_0^2$ то значення величини χ^2 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 $P(\chi^2 < \chi_0^2) = 1 - \varepsilon$, яка нами задається наперед, наша гіпотеза підтверджується. З ймовірністю $P(\chi^2 > \chi_0^2) = \varepsilon$ наша гіпотеза не підтверджується. Звичайно ймовірність ε беруть надзвичайно малою, так щоб спостереження значень випадкової величини, що не узгоджуються з гіпотезою, майже не зустрічались.

Задача 44. При вимірюванні випадкової величини ξ були проведені дві серії випробувань по 500 випробувань у кожній. Результати цих випробувань наведені у наступній таблиці.

вибірка	Інтервали і частоти											
	0	1	2	3	4	5	6	7	8	9	10	11
1	41	34	54	39	49	45	41	33	37	41	47	39
2	36	47	41	47	49	45	32	37	40	41	37	48
3	77	81	95	86	98	90	73	70	77	82	84	87

Значення випадкової величини лежать в інтервалі $[0, 12]$. Цей інтервал розбитий на дванадцять менших інтервалів довжиною 1. Для кожної з двох серій випробувань обчислені частоти попадання значень випадкової величини в той, чи інший інтервал, тобто кількість значень випадкової величини v_i , $i = 0, 1, \dots, 11$, що потрапляють у відповідний інтервал. Останній рядок таблиці містить третю серію випробувань, якою є об'єднання зазначених двох серій випробувань. Відповідні частоти є просто сумами частот перших двох серій випробувань. Гіпотезою є припущення, що дана випадкова величина описується рівномірним розподілом ймовірності. Це означає, що теоретична ймовірність попадання результатів окремих випробувань у будь-який інтервал з дванадцяти дорівнює $p_i = 1/12$, $i = 0, 1, \dots, 11$.

Розв'язання. За міру D відхилення експериментальної функції розподілу ймовірності від теоретичної візьмемо величину

$$\chi^2 = \sum_{i=1}^r \frac{(v_i - np_i)^2}{np_i}.$$

В нашому випадку $n = 500$ для двох серій випробувань і $n = 1000$ для об'єднаної серії випробувань. В усіх випадках $r = 11$. Для першої серії випробувань $\chi_1^2 = 10.000$. Для другої серії випробувань

$\chi^2_2 = 8.032$. Для третьої, об'єднаної, серії випробувань $\chi^2_3 = 9.464$

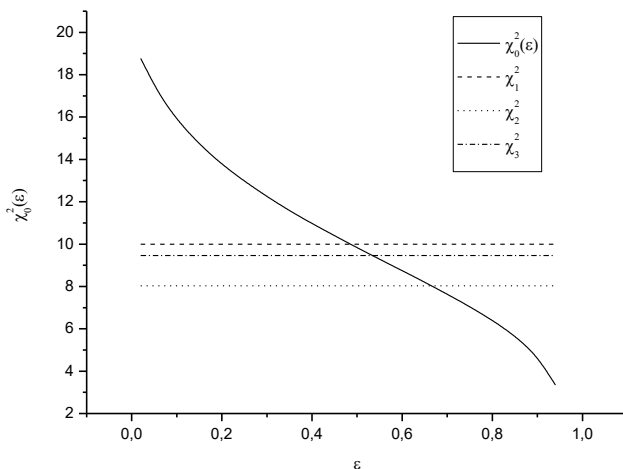
Величина χ^2 є також випадковою і виникає питання про її функцію розподілу. Оскільки обсяг вибірки досить високий, то відповідно до теореми цією функцією розподілу є функція розподілу χ^2

$$P(\chi^2 < \chi^2_0) = 1 - \varepsilon = \int_0^{\chi^2_0} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx .$$

Остання формула визначає ймовірність того, що значення випадкової величини χ^2 є меншими за величину χ^2_0 . Ймовірність протилежної події дорівнює

$$P(\chi^2 \geq \chi^2_0) = \varepsilon = \int_{\chi^2_0}^{\infty} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx .$$

Очевидно, що χ^2_0 є функцією від ε . Тут ε є рівнем значущості, $\chi^2_0(\varepsilon)$ - межею значущості.



Мал. 29. Залежність межі значущості від рівня значущості.

З малюнка видно, що для малих рівнів значущості, наприклад $\varepsilon \approx 0.05$ межа значущості майже вдвічі перевищує отримані значення мір відхилення експериментальної функції розподілу ймовірності від теоретичної χ_1^2 , χ_2^2 , χ_3^2 . Тобто з дуже високою ймовірністю гіпотеза має бути прийнята.

Значно частіше гіпотеза щодо функції розподілу ймовірностей відповідає нашому уявленню про цю функцію з точністю до невідомих її параметрів. Наприклад, ми можемо припустити, що теоретичним розподілом є нормальний розподіл але з невідомими математичним очікуванням і дисперсією. В цьому разі також використовують розподіл χ^2 для оцінки вірогідності гіпотези. Відміна від попереднього випадку - цілком відомої функції розподілу - це менша кількість ступенів вільності у розподілу χ^2 . Це зменшення в точності дорівнює кількості невідомих параметрів розподілу. В якості значень цих параметрів,

необхідних для обчислення теоретичних ймовірностей, беруться їх вибіркві середні.

Задача 45. Розглянемо умови попередньої задачі але в якості гіпотези припустимо, що теоретичний розподіл є нормальним з невідомими параметрами цього розподілу.

Розв'язання. Для параметрів нормального розподілу візьмемо їх вибіркві значення

$$\bar{x} = \frac{1}{n} \sum_{i=1}^r n_i x_i ,$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^r (x_i - \bar{x})^2 n_i ,$$

де в якості значень випадкової величини ξ візьмемо координати точок, що відповідають серединам інтервалів. Тоді $\bar{x} = 5.942$, $S^2 = 11.83$. Нормальний розподіл має вигляд

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

За міру D відхилення експериментальної функції розподілу ймовірності від теоретичної візьмемо величину

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i} ,$$

де теоретичні ймовірності знайдемо для нормального розподілу

$$p_i = \int_{a_i}^{a_{i+1}} P_{\xi}(x) dx = \frac{1}{S\sqrt{2\pi}} \int_{a_i}^{a_{i+1}} \exp\left(-\frac{(x - \bar{x})^2}{2S^2}\right) dx ,$$

$$a = (a_0, a_1, \dots, a_{11}) = (0, 1, \dots, 12) .$$

Обмежимося однією - першою серією випробувань. Тоді $n = 500$, $r = 11$. Для цієї серії випробувань $\chi^2 = 122.094$. Величина χ^2 є також випадковою і виникає питання про її функцію розподілу. Оскільки обсяг вибірки досить високий, то відповідно до теореми

Пірсона, цією функцією розподілу є функція розподілу χ^2 . Врахуємо, що теоретична функція розподілу залежить від двох параметрів, значення яких ми отримали на основі нашої вибірки. Тому для визначення межі значущості слід взяти розподіл χ^2 з $12-1-2=9$ ступенями вільності

$$P(\chi^2 < \chi_0^2) = 1 - \varepsilon = \int_0^{\chi_0^2} \frac{1}{\Gamma\left(\frac{n-1-2}{2}\right) 2^{\frac{n-1-2}{2}}} x^{\frac{n-1-2}{2}-1} e^{-\frac{x}{2}} dx.$$

Остання формула визначає ймовірність того, що значення випадкової величини χ^2 є меншими за величину χ_0^2 . Ймовірність протилежної події дорівнює

$$P(\chi^2 \geq \chi_0^2) = \varepsilon = \int_{\chi_0^2}^{\infty} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx.$$

Легко переконатись, що межа значущості не перевищує 17. Отже, для будь-якого рівня значущості $\chi^2 \gg \chi_0^2$. Таким чином, гіпотеза про нормальний теоретичний розподіл ймовірності не узгоджується з результатами нашої вибірки і отже має бути відкинута. В цьому немає нічого дивного, оскільки в попередній задачі гіпотеза про рівномірний розподіл цієї ж вибірки переконливо підтвердилась.

3.11. Статистична перевірка гіпотези про математичне очікування нормально розподіленої генеральної сукупності

Якщо дисперсія сукупності відома то в якості основної гіпотези H_0 береться припущення про те, що значення

вимірюваної величини (математичного очікування) дорівнює $a = a_0$. В якості альтернативної гіпотези H_1 береться припущення про те, що вимірювана величина (математичне очікування) дорівнює $a = a_1$. Для з'ясування вірності тої, чи іншої гіпотези, слід порівняти обчислене на основі вибірки значення зручної вибіркової функції і значення цієї вибіркової функції, обчислене теоретично на основі припущення про відомий розподіл вимірюваної випадкової величини. Таку зручну вибірку функцію називають **статистичною характеристикою** для перевірки даної гіпотези. У разі нормального розподілу вимірюваної випадкової величини з її відомою дисперсією таку статистичну характеристику легко створити у такому вигляді

$$z = \frac{\bar{x} - a_0}{\sigma} \sqrt{n}.$$

Тут $\bar{x}$ - вибіркове середнє (середнє арифметичне значень елементів вибірки), σ - відоме середнє квадратичне відхилення, n - обсяг вибірки. Оскільки закон розподілу вимірюваної випадкової величини є нормальним, то така статистична характеристика доцільна при довільному обсязі вибірки. Таку ж статистичну характеристику можна взяти і у разі невідомого закону розподілу вимірюваної випадкової величини. Але у цьому разі слід припустити скінченність її дисперсії і використати вибірку досить великого обсягу, щоби виконувались умови Центральної граничної теореми. При цьому в якості дисперсії можна взяти вибірку дисперсію. В обох зазначених випадках функція розподілу ймовірностей матиме вигляд

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

де $P_{\xi}(x)$ густина розподілу. Статистична характеристика у нашому випадку підібрана так, щоби вона мала нормальний

розподіл з нульовим математичним очікування і одиничною дисперсією

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

Очевидно, що ймовірність того, що значення вибіркового середнього, яке наближено уособлює значення вимірюваної величини не перевищує певне критичне значення z_0 дорівнює

$$P(x < z_0) = F_{\xi}(z_0) = \int_{-\infty}^{z_0} P_{\xi}(x) dx.$$

Якщо $a_1 > a_0$, то для справедливості гіпотези H_0 має виконуватись умова

$$a_1 > z_0 > a_0.$$

Тоді з імовірністю $P_{H_0} = F_{\xi}(z_0)$ гіпотезі H_0 слід довіряти, а з імовірністю $P_{\overline{H_0}} = 1 - F_{\xi}(z_0)$ її слід відкинути. Якщо виявилось, що $a_0 < z_0$, то гіпотезу H_0 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{\overline{H_0}} = 1 - F_{\xi}(z_0)$. Величина

$$\alpha = P_{\overline{H_0}} = 1 - F_{\xi}(z_0)$$

називається **рівнем значущості**. Якщо ж виявилось, що $a_0 \geq z_0$, то гіпотезу H_0 слід відхилити на користь гіпотези H_1 . Якщо, насправді, гіпотеза H_0 все таки вірна, то ми при цьому робимо помилку, яка називається **помилкою першого роду**. Область $[z_0, \infty)$ ще називається критичною областю для гіпотези H_0 . Така критична область називається правосторонньою.

Очевидно, що так само можна перевіряти і гіпотезу H_1 . Тоді з імовірністю $P_{H_1} = P_{\overline{H_0}} = 1 - F_{\xi}(z_0)$ гіпотезі H_1 слід довіряти, а з імовірністю $P_{\overline{H_1}} = P_{H_0} = F_{\xi}(z_0)$ її слід відкинути. Якщо

виявилось, що $a_1 > z_0$, то гіпотезу H_1 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{\overline{H_1}} = F_{\xi}(z_0)$.

Якщо ж виявилось, що $a_0 \leq z_0$, то гіпотезу H_1 слід відхилити на користь гіпотези H_0 . Якщо, насправді, гіпотеза H_1 все таки вірна, то ми при цьому робимо помилку, яка називається **помилкою другого роду**. Область $(-\infty, z_0)$ ще називається критичною областю для гіпотези H_1 .

Порівнюючи ймовірності помилок першого і другого родів, можна побачити, що перша з них зменшується в міру наближення z_0 до a_1 . При цьому друга з них відповідно зростає. В міру наближення z_0 до a_0 зростає помилка першого роду, зате спадає помилка другого роду.

Величина $1 - P_{\overline{H_1}} = 1 - F_{\xi}(z_0)$ називається **потужністю критерію**. Задаючи потужність критерію, ми тим самим задаємо критичне значення z_0 , що впливає на величини помилок першого і другого родів.

Якщо $a_1 < a_0$, то для справедливості гіпотези H_0 має виконуватись умова

$$a_1 < z_0 < a_0.$$

Тоді з імовірністю $P_{H_0} = 1 - F_{\xi}(z_0)$ гіпотезі H_0 слід довіряти, а з імовірністю $P_{\overline{H_0}} = F_{\xi}(z_0)$ її слід відкинути. Якщо виявилось, що $a_0 > z_0$, то гіпотезу H_0 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{\overline{H_0}} = F_{\xi}(z_0)$. Якщо ж виявилось, що $a_0 \leq z_0$, то гіпотезу H_0 слід відхилити на користь гіпотези H_1 . Якщо, насправді, гіпотеза H_0 все таки вірна, то ми при цьому робимо помилку, яка називається помилкою першого

роду. Область $(-\infty, z_0]$ знову ж таки називається критичною областю для гіпотези H_0 . Така критична область називається лівосторонньою.

Очевидно, що так само можна перевіряти і гіпотезу H_1 . Тоді з імовірністю $P_{H_1} = P_{\overline{H_0}} = F_\xi(z_0)$ гіпотезі H_1 слід довіряти, а з імовірністю $P_{\overline{H_1}} = P_{H_0} = 1 - F_\xi(z_0)$ її слід відкинути. Якщо виявилось, що $a_1 < z_0$, то гіпотезу H_1 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{\overline{H_1}} = 1 - F_\xi(z_0)$. Якщо ж виявилось, що $a_1 \geq z_0$, то гіпотезу H_1 слід відхилити на користь гіпотези H_0 . Якщо, насправді, гіпотеза H_1 все таки вірна, то ми при цьому робимо помилку, яка називається помилкою другого роду. Область $[z_0, \infty)$ ще називається критичною областю для гіпотези H_1 .

Порівнюючи ймовірності помилок першого і другого родів, можна побачити, що перша з них зменшується в міру наближення z_0 до a_1 . При цьому друга з них відповідно зменшується. В міру наближення z_0 до a_0 зростає помилка першого роду, зате спадає помилка другого роду.

Величина $1 - P_{\overline{H_1}} = F_\xi(z_0)$ називається потужністю критерію у цьому разі.

Виникає необхідність і у перевірці гіпотези, про рівність значень двох вимірюваних випадкових величин, тобто

$$a_0 = a_1.$$

У цьому разі гіпотеза H_0 про рівність двох вимірюваних величин буде прийматись, якщо їх вибіркові середні належатимуть інтервалу $(-z_0, z_0)$, де z_0 критичне значення статистичної характеристики. Ймовірність справедливості цієї гіпотези

$$P_{H_0}(-z_0 < \bar{x}_{a_0} < z_0) = F_{\xi}(z_0) = \int_{-z_0}^{z_0} P_{\xi}(x) dx$$

або

$$P_{H_0}(-z_0 < \bar{x}_{a_1} < z_0) = F_{\xi}(z_0) = \int_{-z_0}^{z_0} P_{\xi}(x) dx ,$$

причому нерівності $-z_0 < \bar{x}_{a_0} < z_0$ і $-z_0 < \bar{x}_{a_1} < z_0$ виконуються одночасно. Якщо хоча б одна з цих нерівностей не виконується, то гіпотеза H_0 відкидається на користь гіпотези H_1 про те, що дані величини мають різні значення. Критична область гіпотези H_0 визначається так $(-\infty, -z_0] \cup [z_0, \infty)$. Очевидно, що чим більше ми звужуємо інтервал $[-z_0, z_0]$, тим меншою є ймовірність прийняття гіпотези H_0 .

Якщо дисперсія генеральної сукупності невідома, то в якості статичної характеристики доцільно взяти функцію

$$z = \frac{\bar{x} - a_0}{S} \sqrt{n-1} ,$$

розподілену за законом Ст'юдента з $n-1$ ступенями вільності. Тут S^2 - виправлена вибіркова дисперсія. Густина розподілу Ст'юдента має вигляд

$$P_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} , \quad \tau_n = \frac{\xi_0}{\sqrt{\frac{1}{n} \chi_n^2}} , \quad x \in (-\infty, \infty) .$$

Відповідно для функції розподілу матимемо

$$F_{\tau_n}(x) = \int_{-\infty}^x P_{\tau_n}(x) dx$$

3.12. Статистична перевірка гіпотези про дисперсію нормально розподіленої генеральної сукупності

Порівнюючи два значення дисперсії випадкової величини σ_0^2 і σ_1^2 для з'ясування того, яке з них вірне, дотримуватимемось тої ж схеми, що і у разі математичного очікування. Гіпотезою H_0 вважатимемо припущення про те, що вірним є значення σ_0^2 . Альтернативною гіпотезою H_1 вважатимемо припущення про те, що вірним є значення σ_1^2 . Зручною статистичною характеристикою, очевидно, є наступна вибіркова функція

$$z = \frac{(n-1)S^2}{\sigma_0^2},$$

де S^2 є виправлена вибірковою дисперсією. Відомо, що наведена вибіркова функція має розподіл χ^2 з $n-1$ ступенями вільності. Функція розподілу χ^2 має вигляд

$$F_{\chi_n^2}(x) = \int_0^x P_{\chi_n^2}(x) dx$$

У разі нормально розподілених випадкових величин $\xi_0, \xi_1, \dots, \xi_n$ з параметрами $(0,1)$ густина цього розподілу така

$$P_{\chi_n^2}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}},$$

де

$$\chi_n^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2.$$

Очевидно, що ймовірність того, що значення випадкової величини $\chi_n^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2$ не перевищує певне критичне значення z_0 дорівнює

$$P(\chi < z_0) = F_{\chi_{n-1}^2}(z_0) = \int_0^{z_0} P_{\chi_{n-1}^2}(x) dx.$$

Якщо $\sigma_1^2 > \sigma_0^2$, то для справедливості гіпотези H_0 має виконуватись умова

$$\sigma_1^2 > z_0 / (n-1) > \sigma_0^2.$$

Тоді з імовірністю $P_{H_0} = F_{\chi_{n-1}^2}(z_0)$ гіпотезі H_0 слід довіряти, а з імовірністю $P_{\bar{H}_0} = 1 - F_{\chi_{n-1}^2}(z_0)$ її слід відкинути. Якщо виявилось, що $(n-1)S_0^2 < z_0$, то гіпотезу H_0 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{\bar{H}_0} = 1 - F_{\chi_{n-1}^2}(z_0)$. Величина

$$\alpha = P_{\bar{H}_0} = 1 - F_{\chi_{n-1}^2}(z_0)$$

називається рівнем значущості. Якщо ж виявилось, що $(n-1)S_0^2 \geq z_0$, то гіпотезу H_0 слід відхилити на користь гіпотези H_1 . Якщо, насправді, гіпотеза H_0 все таки вірна, то ми при цьому робимо помилку, яка називається помилкою першого роду. Область $[z_0, \infty)$ ще називається критичною областю для гіпотези H_0 . Така критична область називається правосторонньою.

Очевидно, що так само можна перевіряти і гіпотезу H_1 . Тоді з імовірністю $P_{H_1} = P_{\bar{H}_0} = 1 - F_{\chi_{n-1}^2}(z_0)$ гіпотезі H_1 слід довіряти, а з імовірністю $P_{\bar{H}_1} = P_{H_0} = F_{\chi_{n-1}^2}(z_0)$ її слід відкинути. Якщо виявилось, що $(n-1)S_1^2 > z_0$, то гіпотезу H_1 слід прийняти.

При цьому ймовірність помилковості прийнятого рішення буде $P_{H_1} = F_{\chi_{n-1}^2}(z_0)$. Якщо ж виявилось, що $(n-1)S_1^2 \leq z_0$, то гіпотезу H_1 слід відхилити на користь гіпотези H_0 . Якщо, насправді, гіпотеза H_1 все таки вірна, то ми при цьому робимо помилку, яка називається помилкою другого роду. Область $(0, z_0)$ ще називається критичною областю для гіпотези H_1 .

Порівнюючи ймовірності помилок першого і другого родів, можна побачити, що перша з них зменшується в міру наближення z_0 до $(n-1)S_1^2$. При цьому друга з них відповідно зменшується. В міру наближення z_0 до $(n-1)S_0^2$ зростає помилка першого роду, зате спадає помилка другого роду.

Величина $1 - P_{H_1} = 1 - F_{\chi_{n-1}^2}(z_0)$ називається потужністю критерію. Задаючи потужність критерію, ми тим самим задаємо критичне значення z_0 , що впливає на величини помилок першого і другого родів.

Якщо $S_1^2 < S_0^2$, то для справедливості гіпотези H_0 має виконуватись умова

$$S_1^2 < z_0 / (n-1) < S_0^2.$$

Тоді з ймовірністю $P_{H_0} = 1 - F_{\chi_{n-1}^2}(z_0)$ гіпотезі H_0 слід довіряти, а з ймовірністю $P_{H_0} = F_{\chi_{n-1}^2}(z_0)$ її слід відкинути. Якщо виявилось, що $(n-1)S_0^2 > z_0$, то гіпотезу H_0 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{H_0} = F_{\chi_{n-1}^2}(z_0)$. Якщо ж виявилось, що $(n-1)S_0^2 \leq z_0$, то гіпотезу H_0 слід відхилити на користь гіпотези H_1 . Якщо, насправді,

гіпотеза H_0 все таки вірна, то ми при цьому робимо помилку, яка називається помилкою першого роду. Область $(0, z_0]$ знову ж таки називається критичною областю для гіпотези H_0 . Така критична область називається лівосторонньою.

Очевидно, що так само можна перевіряти і гіпотезу H_1 . Тоді з імовірністю $P_{H_1} = P_{\overline{H_0}} = F_{\chi_{n-1}^2}(z_0)$ гіпотезі H_1 слід довіряти, а з імовірністю $P_{\overline{H_1}} = P_{H_0} = 1 - F_{\chi_{n-1}^2}(z_0)$ її слід відкинути. Якщо виявилось, що $(n-1)S_1^2 < z_0$, то гіпотезу H_1 слід прийняти. При цьому ймовірність помилковості прийнятого рішення буде $P_{\overline{H_1}} = 1 - F_{\chi_{n-1}^2}(z_0)$. Якщо ж виявилось, що $(n-1)S_1^2 > z_0$, то гіпотезу H_1 слід відхилити на користь гіпотези H_0 . Якщо, насправді, гіпотеза H_1 все таки вірна, то ми при цьому робимо помилку, яка називається помилкою другого роду. Область $[z_0, \infty)$ ще називається критичною областю для гіпотези H_1 .

Порівнюючи ймовірності помилок першого і другого родів, можна побачити, що перша з них зменшується в міру наближення z_0 до $(n-1)S_1^2$. При цьому друга з них відповідно зменшується. В міру наближення z_0 до $(n-1)S_0^2$ зростає помилка першого роду, зате спадає помилка другого роду.

Величина $1 - P_{\overline{H_1}} = F_{\chi_{n-1}^2}(z_0)$ називається потужністю критерію у цьому разі.

Виникає необхідність і у перевірці гіпотези, про рівність значень двох вимірюваних випадкових величин, тобто

$$\sigma_0^2 = \sigma_1^2.$$

У цьому разі гіпотеза H_0 про рівність двох вимірюваних величин буде прийматись, якщо їх вибіркові середні належатимуть інтервалу $(-z_0, z_0)$, де z_0 критичне значення статистичної характеристики. Ймовірність справедливості цієї гіпотези

$$P_{H_0}(-z_0 < S_0^2 < z_0) = F_{\chi_{n-1}^2}(z_0) = \int_{-z_0}^{z_0} P_{\chi_{n-1}^2}(x) dx$$

або

$$P_{H_0}(-z_0 < S_1^2 < z_0) = F_{\chi_{n-1}^2}(z_0) = \int_{-z_0}^{z_0} P_{\chi_{n-1}^2}(x) dx,$$

причому нерівності $-z_0 < (n-1)S_0^2 < z_0$ і $-z_0 < (n-1)S_1^2 < z_0$ виконуються одночасно. Якщо хоча б одна з цих нерівностей не виконується, то гіпотеза H_0 відкидається на користь гіпотези H_1 про те, що дані величини мають різні значення. Критична область гіпотези H_0 визначається так $(0, -z_0] \cup [z_0, \infty)$. Очевидно, що чим більше ми звужуємо інтервал $[-z_0, z_0]$, тим меншою є ймовірність прийняття гіпотези H_0 .

3.13. Регресійний аналіз

До задач оцінки параметрів часто приводяться задачі, у яких потрібно встановити залежність між змінними. Наприклад, нехай відомо, що змінна y лінійно залежить від змінних $x_1, x_2, \dots, x_k$, тобто $y = A_0 + A_1x_1 + \dots + A_kx_k$, де коефіцієнти $A_0, A_1, \dots, A_k$ невідомі. Для їх знаходження використовуються результати ряду вимірювань величин $y, x_1, x_2, \dots, x_k$. Нехай таких вимірювань n і у кожному з них отриманий певний набір значень вимірюваних величин $y_i, x_{i1}, x_{i2}, \dots, x_{in}$, $i = 1, \dots, n$. За результатами

цих вимірювань можна знайти невідомі коефіцієнти $A_0, A_1, \dots, A_n$. Наведена задача відноситься до регресійних задач.

Регресійні задачі відрізняються від задач інтерполяційних. В останніх також за результатами вимірювань будується функція, як правило поліном, така, що їх значення в усіх точках $U_i, X_{i1}, X_{i2}, \dots, X_{in}$, 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 $U_i, X_{i1}, X_{i2}, \dots, X_{in}$ 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 вимірних значень $U_i, X_{i1}, X_{i2}, \dots, X_{in}$ – 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, b лінійної функції $y = ax + b$. Принциповим тут є те, що для кожного конкретного x величина y вимірюється із скінченною похибкою δ .

Для визначення сталих a і b проведемо n вимірювань величини y . Нехай це будуть значення $y_1, y_2, \dots, y_n$, для відповідного набору значень величини x : $x_1, x_2, \dots, x_n$. При цьому нехай значення $x_1, x_2, \dots, x_n$ знайдені з похибкою δi , $i = 1, \dots, n$.

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

$$P(\delta i, a, b, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\delta i^2}{2\sigma^2}\right),$$

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

$$P(y_i, a, b, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y_i - ax_i - b)^2}{2\sigma^2}\right].$$

Цей розподіл зручніший, оскільки в нього входять лише вимірювані величини x_i, y_i . Функція правдоподібності має вигляд

$$L = L(y_1, \dots, y_n, a, b, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left[-\frac{1}{2\sigma^2} \sum_{k=1}^n (y_k - ax_k - b)^2\right].$$

Тепер легко отримати і систему рівнянь правдоподібності

$$\begin{aligned}\frac{\partial \ln(L)}{\partial a} &= \frac{1}{\sigma^2} \sum_{k=1}^n x_i (y_i - ax_i - b) = 0, \\ \frac{\partial \ln(L)}{\partial b} &= \frac{1}{\sigma^2} \sum_{k=1}^n (y_i - ax_i - b) = 0, \\ \frac{\partial \ln(L)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{k=1}^n (y_i - ax_i - b)^2 = 0.\end{aligned}$$

Розв'язання цієї системи рівнянь не становить жодних проблем, проте найпростіший випадок виникає для $x_1 + x_2 + \dots + x_n = 0$. Тоді

$$\begin{aligned}\tilde{a} &= \sum_{k=1}^n x_k y_k / \sum_{k=1}^n x_k^2, \\ \tilde{b} &= \frac{1}{n} \sum_{k=1}^n y_k, \\ \tilde{\sigma}^2 &= \frac{1}{n} \sum_{k=1}^n (y_i - \tilde{a}x_i - \tilde{b})^2.\end{aligned}$$

Цим формулам можна надати і іншого вигляду, виключивши з них y_i . Тепер

$$\begin{aligned}\tilde{a} &= a + \sum_{k=1}^n x_i \delta_i / \sum_{k=1}^n x_i^2, \\ \tilde{b} &= b + \frac{1}{n} \sum_{k=1}^n \delta_i.\end{aligned}$$

Останні дві рівності дозволяють легко перевірити незміщеність і слушність наших оцінок a і b . Дійсно, незміщеність впливає з того, що

$$M\tilde{a} = a, \quad M\tilde{b} = b,$$

Слушність з того, що

$$D\tilde{a} = \sigma^2 / \sum_{k=1}^n x_k^2, \quad D\tilde{b} = \sigma^2 / n,$$

а тому

$$\lim_{n \rightarrow \infty} D\tilde{a} = \lim_{n \rightarrow \infty} D\tilde{b} = 0 .$$

Якщо припущення про нормальність розподілу δi не відповідає дійсності, то оцінки параметрів a і b можна отримати з умови мінімуму виразу

$$Q(a, b) = \sum_{k=1}^n (y_k - ax_k - b)^2 = \sum_{k=1}^n \delta_k^2 ,$$

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 $Q(a, b)$

$$\frac{\partial Q(a, b)}{\partial a} = 2 \sum_{k=1}^n x_k (y_k - ax_k - b) = 0 ,$$

$$\frac{\partial Q(a, b)}{\partial b} = 2 \sum_{k=1}^n (y_k - ax_k - b) = 0 .$$

Отримані рівняння збігаються з відповідними рівняннями методу найбільшої правдоподібності і у разі, якщо $x_1 + x_2 + \dots + x_n = 0$ дають вже наведений вище результат

$$\tilde{a} = \sum_{k=1}^n x_k y_k / \sum_{k=1}^n x_k^2 ,$$

$$\tilde{b} = \frac{1}{n} \sum_{k=1}^n y_k .$$

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 δ_i 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.

3.14. 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

$$y = f(x; a_0, a_1, \dots, a_n),$$

де $a_0, a_1, \dots, a_n$ – 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

$$S = \sum_{k=1}^N [y_k - f(x_k; a_0, a_1, \dots, a_n)]^2$$

- сума квадратів відхилень вимірюваної фізичної величини у від розрахованих $f(x; a_0, a_1, \dots, a_n)$ приймає мінімальне значення. Наведена формула вірна для набору експериментальних даних, вимірюваних з однаковою точністю (дисперсією). Якщо всі вимірювані результати мають різну точність (дисперсію σ_i), то остання сума заміниться наступною

$$S = \sum_{k=1}^N [y_k - f(x_k; a_0, a_1, \dots, a_n)] / \sigma_k^2.$$

З умови екстремуму цієї функції ми і отримаємо систему рівнянь для знаходження невідомих параметрів

$$\frac{\partial S(a_0, a_1, \dots, a_n)}{\partial a_0} = 0,$$

$$\frac{\partial S(a_0, a_1, \dots, a_n)}{\partial a_1} = 0,$$

.....

$$\frac{\partial S(a_0, a_1, \dots, a_n)}{\partial a_n} = 0.$$

Ясно, що в загальному випадку цю систему рівнянь можна розв'язати лише у числовий спосіб.

3.15. Дисперсійний аналіз

Дисперсійний аналіз стосується наступного класу задач. Нехай одна й та сама величина вимірюється за допомогою різних приладів. Нас цікавить, чи мають різні прилади систематичну похибку. Очевидно процес вимірювання доцільно провести так. Відповідно до кількості приладів r потрібно провести r серій вимірювань. У кожній i -ій серії нехай виконано n_i вимірювань $x_{i1}, x_{i2}, \dots, x_{in_i}$. Як завжди, результати окремих вимірювань вважатимемо незалежними випадковими величинами. Для простоти вважатимемо, що всі вони мають нормальний закон розподілу з однаковими для кожної серії (кожного приладу) математичним очікуванням $Mx_{ij} = a_i$, $i = 1, 2, \dots, r$. Наявність чи відсутність систематичної похибки будемо відслідковувати за допомогою математичного очікування. Тобто перевіримо, чи $a_1 = a_2 = \dots = a_r = a$. При цьому також, для простоти, вважатимемо дисперсію всіх випадкових величин для всіх серій однаковою,

тобто $Dx_{ij} = \sigma^2$, $i = 1, 2, \dots, r$, $j = 1, 2, \dots, n_i$. Введемо далі вибіркове середнє для окремих серій

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij},$$

та загальне вибіркове середнє для всіх серій

$$\bar{x} = \frac{1}{n} \sum_{i=1}^r \sum_{j=1}^{n_i} x_{ij},$$

де

$$n = n_1 + \dots + n_r.$$

Легко перевірити, що введені нами групові вибірккові середні є незміщеними і слушними оцінками величин a_i . Якщо всі a_i однакові, то загальне вибіркове середнє не повинне суттєво відрізнятись від групових. Натомість різниця між $\bar{x}_i$ і $\bar{x}$ має бути суттєвою. Розглянемо загальну суму квадратів відхилень значень випадкових величин від загально вибіркового середнього

$$Q = \sum_{i=1}^r \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2.$$

Цю функцію зручно розбити на два доданки

$$Q = Q_1 + Q_2,$$

де перший з них Q_1 є сумою квадратів відхилень групових середніх від загального середнього з ваговими множниками, рівними кількості елементів у групах

$$Q_1 = \sum_{i=1}^r n_i (\bar{x}_i - \bar{x})^2.$$

Другий доданок є сумою квадратів відхилень випадкових величин однієї групи від групового вибіркового середнього, підсумована по всім групам

$$Q_2 = \sum_{i=1}^r \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2.$$

Суму Q_1 ще називають сумою квадратів відхилень між групами, Q – сумою квадратів відхилень всередині групи. Доведено, що при умові $a_1 = a_2 = \dots = a_r = a$ випадкові величини Q_1 і Q_2 незалежні.

Отже ми маємо дві суми квадратів незалежних, нормально розподілених випадкових величин з відомою дисперсією. Відповідно до загальної теорії кожна з цих сум квадратів розподілена за законом χ^2 . Детальніше це виглядає так.

Випадкова величина $\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$ розподілена за законом χ^2 з

$n_i - 1$ степенями свободи. Відповідно випадкова величина Q_2 розподілена також за законом χ^2 , але з $\sum_{i=1}^r (n_i - 1) = n - r$

степенями свободи. Незалежна від неї випадкова величина Q_1 також розподілена за законом χ^2 з $r - 1$ степенями свободи. Наведених міркувань достатньо, щоб стверджувати, що

$$F_{r-1, n-r} = \frac{Q_1 / (r-1)}{Q_2 / (n-r)}$$

має розподіл Фішера з $r - 1$ і $n - r$ степенями свободи.

Статистичний критерій формулюється як звичайна інтервальна оцінка. При заданому рівні значимості α

$$P[F_{r-1, n-r} > C(\alpha)] = \alpha$$

- ймовірність того, що прилади мають систематичну похибку,

$$P[F_{r-1, n-r} < C(\alpha)] = 1 - \alpha$$

- ймовірність того, що прибори такої похибки не мають. Часто рівень значимості беруть рівним $\alpha = 0.05$, тоді для $n^2 = 10$, $n_i = 10$, $C(0.05) = 2.97$. Отже, якщо ми обчислили $F_{r-1, n-r}$ і виявилось, що $F_{r-1, n-r} > C(\alpha)$, то відбулась малоймовірна подія, що ставить під сумнів умову $a_1 = a_2 = \dots = a_r = a$. Якщо ж $F_{r-1, n-r} < C(\alpha)$, то відбулась дуже ймовірна подія і припущення має сенс з ймовірністю $1 - \alpha$ і з ймовірністю α .

3. 16. Задачі для самостійного розв'язання

Варіант 1.

1. Маємо 4 заготовки для виготовлення деталей. Ймовірність виготовлення придатної деталі дорівнює 0.75. Знайти закон розподілу випадкової величини, що є кількістю заготовок, які буде використано для виготовлення придатної деталі. Знайти математичне очікування і дисперсію випадкової величини, а також імовірність того, що із цих заготовок буде виготовлена 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 оцінку для параметра σ^2 і перевірити її на незміщеність, обґрунтованість і ефективність.

5. Середнє споживання електричної енергії, обчислене протягом багатьох років, становить на рік 300000 кВт год.

1) Оцінити ймовірність того, що споживання електричної енергії протягом наступного року перевищить 500000 кВт год.

2) Оцінити ту саму ймовірність за умови, що середнє квадратичне відхилення споживання енергії за рік від його середнього значення становить 50000 кВт.

6. За даними вибірки обсягом n з нормально розподіленої генеральної сукупності

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

дисперсія якої відома, а довірча ймовірність дорівнює γ , знайти довірчий інтервал для математичного очікування.

Варіант 2.

1. Задано функцію

$$F_{\xi}(x) = \begin{cases} 0, & \text{if } x \leq 1; \\ ax^2 + bx, & \text{if } 1 < x \leq 4; \\ 1, & \text{if } x > 3. \end{cases}$$

Довести, що можна дібрати такі значення a і b , при яких функція $F_{\xi}(x)$ буде функцією розподілу ймовірностей випадкової величини $F_{\xi}(x)$ буде функцією розподілу випадкової величини ξ . Знайти $P(2 \leq \xi \leq 3)$.

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 інтервальный ряд і побудувати гістограму. Запропонувати гіпотезу про вигляд

функції розподілу ймовірності для сукупності. Знайти вибіркові середнє, дисперсію, асиметрію і ексцес.

4. За методом моментів знайти оцінку параметра p геометричного розподілу

$$P(m) = p(1-p)^{m-1}, \quad m \in N$$

за даними вибірки обсягом n .

5. Ймовірність деякої події визначається експериментально. Оцінити кількість незалежних випробувань, які забезпечують з імовірністю не менше як 0.99 обчислення шуканої ймовірності з похибкою, що не перевищує 0.01. Оцінку побудувати за допомогою нерівності Чебишева і інтегральної теореми Муавра-Лапласа.

6. За даними вибірки обсягом n з нормально розподіленої генеральної сукупності

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

дисперсія якої невідома, а довірча ймовірність дорівнює γ , знайти довірчий інтервал для математичного очікування.

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 функції розподілу і знайти $|F_{\xi}(x) - F_{\xi}^*(x)|$, вважаючи, що кількість використовуваних деталей має розподіл Пуассона з

параметром розподілу, рівним 1. Обчислити вибіркowi середнє і дисперсію і співставити їх з математичним очікуванням і дисперсією, згідно з висунутою гіпотезою про закон розподілу у сукупності.

4. Із нормально розподіленої сукупності

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

з $D\xi = \sigma^2$ зроблено вибірку обсягом n . Знайти оцінку для $M\xi$. Перевірити її на незміщеність ефективність та обґрунтованість.

5. Середні витрати води, обчислені протягом багатьох місяців, становлять на місяць 30000 л.

1) Оцінити ймовірність того, що споживання води протягом наступного місяця перевищить 50000 л.

2) Оцінити ту саму ймовірність за умови, що середнє квадратичне відхилення споживання енергії за рік від його середнього значення становить 5000 л.

6. За даними вибірки обсягом n з нормально розподіленої генеральної сукупності

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

математичне очікування якої відоме, а довірча ймовірність дорівнює γ , знайти довірчий інтервал для дисперсії.

Варіант 4.

1. Випадкова величина задається функцією розподілу

$$F_{\xi}(x) = \begin{cases} 0, & \text{if } x \leq 1; \\ a + b \arctan(x), & \text{if } x > 1. \end{cases}$$

Знайти a , b , $P(2 \leq \xi \leq 4)$ і густину розподілу $P_{\xi}(x)$.

2. Знайти математичне очікування випадкової величини, що має розподіл Бернуллі.

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 і знайти $|F_{\xi}(x) - F_{\xi}^*(x)|$, вважаючи, що справджується геометричний розподіл з $p = 0.25$. Знайти вибіркові середні і дисперсію, а також математичне очікування і дисперсію для геометричного розподілу.

4. Із нормально розподіленої сукупності

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

з $M\xi = a$ зроблено вибірку обсягом n . Знайти оцінку для $D\xi$. Перевірити її на незміщеність ефективність та обґрунтованість.

5. Ймовірність певної події в кожному випробуванні є $1/3$. Яку найменшу кількість випробувань потрібно виконати, щоб з ймовірністю не меншою за 0.99 частота події відхилялась за абсолютною величиною від її ймовірності не більше ніж на 0.001 ? При розв'язанні використати нерівність Чебишева та інтегральну теорему Муавра-Лапласа.

6. За даними вибірки обсягом n з нормально розподіленої генеральної сукупності

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

математичне очікування якої невідоме, а довірча ймовірність дорівнює γ , знайти довірчий інтервал для дисперсії.

Варіант 5.

1. Випадкова величина задається функцією розподілу

$$F_{\xi}(x) = \begin{cases} 0, & \text{if } x \leq a \\ ab^2 - 1/3, & \text{if } a < x \leq 2; \\ 1, & \text{if } x > 2. \end{cases}$$

Знайти a , b , $P(2 \leq \xi \leq 4)$ і густину розподілу $P_{\xi}(x)$.

2. Знайти дисперсію випадкової величини, що має розподіл Коші.

3. У вимірювальному приладі установлена 5 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 $|F_{\xi}(x) - F_{\xi}^*(x)|$, вважаючи, що кількість замінених опорів має розподіл Пуассона з μ рівним 3. Знайти вибіркowi середнє та дисперсію і співставити їхні значення з числовими характеристиками відповідного розподілу Пуассона.

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 . Знайти оцінку для μ і перевірити її на незміщеність, ефективність та обґрунтованість (значення σ^2 відоме).

5. Ймовірність певної події в кожному випробуванні є $1/5$. Яку найменшу кількість випробувань потрібно виконати, щоб з ймовірністю не меншою за 0.97 частота події відхилялась за абсолютною величиною від її ймовірності не більше ніж на 0.005? При розв'язанні використати нерівність Чебишева та інтегральну теорему Муавра-Лапласа.

6. Визначити мінімальний обсяг вибірки для того, щоб із довірчою ймовірністю 0.98 дістати оцінку математичного очікування нормально розподіленої сукупності - різниця вибіркового середнього і математичного очікування менше, ніж 0.2. Середнє квадратичне відхилення для елементів вибірки вважати відомим і рівним 1.5.

Варіант 6.

1. Графік густини розподілу - півколо з центром у початку координат. Знайти аналітичний вираз для густини розподілу, функції 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 інтервальний ряд і побудувати гістограму. Висунути гіпотезу про закон розподілу в сукупності. Знайти вибіркові середнє і дисперсію.

4. Із сукупності із густиною ймовірності

$$P_{\xi}(x) = \begin{cases} a \exp(-a(x - \mu)^2), & \text{if } x \geq \mu, \\ 0, & \text{if } x < \mu \end{cases}$$

зроблено вибірку обсягом n . Знайти оцінку для a і μ .

5. Ймовірність певної події в кожному випробуванні є $1/4$. Яку найменшу кількість випробувань потрібно виконати, щоб з ймовірністю не меншою за 0.98 частота події відхилялась за абсолютною величиною від її ймовірності не більше ніж на 0.002 ? При розв'язанні використати нерівність Чебишева та інтегральну теорему Муавра-Лапласа.

6. 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 з довірчою ймовірністю 0.96 не перевищувала 15 м.

Варіант 7.

1. Випадкова величина набуває значень на відрізку $[a, 2]$, де задано густину розподілу $P_{\xi}(x) = \frac{3}{16} x^{n+1}$. Визначити $a, n (n \in N)$, 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 інтервальний ряд і побудувати гістограму. Висунути гіпотезу про закон розподілу у вибірці. Обчислити вибіркові середнє і дисперсію.

4. Із сукупності із густиною ймовірності

$$P_{\xi}(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right), & \text{if } x > 0, \\ 0, & \text{if } x \leq 0 \end{cases}$$

зроблено вибірку обсягом n . Знайти оцінку для σ^2 і μ .

5. Ймовірність деякої події визначається експериментально. Оцінити кількість незалежних випробувань, які забезпечують з імовірністю не менше як 0.97 обчислення шуканої ймовірності з похибкою, що не перевищує 0.03. Оцінку побудувати за допомогою нерівності Чебишева і інтегральної теореми Муавра-Лапласа.

6. У результаті вимірювання ємності 20 конденсаторів дістали для вибіркового середнього значення 4.47 пф, а для вибіркової дисперсії 0.0121 пф². Порівняйте точність оцінки математичного очікування при довірчій ймовірності 0.95 використовуючи нормальний закон розподілу і розподіл Ст'юдента.

Варіант 8.

1. Випадкова величина набуває 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 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 оцінку для параметра α і показати, що вона має додатне зміщення. Усунувши зміщення, переконатись в тому, що така оцінка ефективна.

5. Середні витрати води, обчислені протягом багатьох місяців, становлять на місяць 40000 л.

1) Оцінити ймовірність того, що споживання води протягом наступного місяця перевищить 60000 л.

2) Оцінити ту саму ймовірність за умови, що середнє квадратичне відхилення споживання енергії за рік від його середнього значення становить 4000 л.

6. Під час перевірки 400 лампочок середній термін їх горіння становив 1220 год. Оцінити з довірчою ймовірністю 0.95 математичне очікування тривалості горіння, якщо середнє квадратичне відхилення дорівнює 35 год., а закон розподілу вибіркового середнього можна вважати нормальним.

Варіант 9.

1. Випадкова величина набуває значень на відрізку $[a, b]$, де задано густину розподілу $P_{\xi}(x) = \frac{3}{16}(x-c)^2$. Визначити a, b , якщо $M\xi = 2$, $P_{\xi}(a) = P_{\xi}(b)$.

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 інтервальний 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

$$P_{\xi}(x) = \lambda \exp(-\lambda x), \quad x \geq 0$$

зроблено вибірку обсягом n . Знайти оцінку для параметра λ , ліквідувати зміщення і перевірити, чи буде оцінка ефективною.

5. Ймовірність деякої події визначається експериментально. Оцінити кількість незалежних випробувань, які забезпечують з імовірністю не менше як 0.98 обчислення шуканої ймовірності з похибкою, що не перевищує 0.02. Оцінку побудувати за допомогою нерівності Чебишева і інтегральної теореми Муавра-Лапласа.

6. 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, b , якщо

$$M\xi = \frac{e^2 + 1}{4}, \quad P_{\xi}(a) = P_{\xi}(b).$$

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 статистичну функції розподілу і знайти $|F_{\xi}(x) - F_{\xi}^*(x)|$, вважаючи, що кількість використовуваних деталей має розподіл Пуассона з параметром розподілу, рівним 1. Обчислити вибіркові середнє і дисперсію і співставити їх з математичним очікуванням і дисперсією, згідно з висунутою гіпотезою про закон розподілу у сукупності.

4. Із нормально розподіленої сукупності

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

з $M\xi = a$ зроблено вибірку обсягом n . Знайти оцінку для $D\xi$.
Перевірити її на незміщеність ефективність та обґрунтованість.

5. Середнє споживання води, обчислене протягом багатьох років, становить на рік 400000 м³.

1) Оцінити ймовірність того, що споживання води протягом наступного року перевищить 600000 м³.

2) Оцінити ту саму ймовірність за умови, що середнє квадратичне відхилення споживання води за рік від його середнього значення становить 40000 м³.

6. Систематична помилка вимірювального приладу дорівнює нулю, а середнє квадратичне відхилення - 20 м. Різниця між вибіркоvim середнім і точним результатом не перевищує 10 м. Вважаючи помилку розподіленою за нормальним законом, визначити довірчу ймовірність для обсягів вибірки 4, 9, 16, 25.

Глава 4. Випадкові процеси

1. Функція розподілу випадкового процесу

Нехай $\xi(t)$ є випадковою функцією часу (випадковий процес). Для кожного фіксованого значення часу $t = t_i$ $\xi_i = \xi(t_i)$ є звичайною випадковою величиною. Отже кожний випадковий процес еквівалентний нескінченній сукупності випадкових величин, оскільки час пробігає всі значення, що належать скінченному інтервалу $t \in [t_a, t_b]$. Ця нескінченність є континуальною, тобто найбільшою з можливих нескінченностей. Замість неї прийнято розглядати найменшу з можливих нескінченностей – злічену, що дорівнює кількості натуральних чисел $t \in \{t_i : i = 1, 2, \dots, n, \dots\}$. З теоретичної точки зору – це певне наближення. З практичної точки зору – це абсолютно прийнятна річ. Така заміна еквівалентна вилученню з множини дійсних чисел множини ірраціональних чисел. Остання процедура ніяк не відбивається на якості представлення різноманітних розрахунків, оскільки ми не здатні жодне ірраціональне число навіть записати у вигляді десяткового дробу, а лише раціональне. Таким чином, без суттєвої втрати загальності розглянемо випадковий процес, як злічену сукупність випадкових величин $\xi(t) = \{\xi_1, \xi_2, \dots, \xi_n, \dots\}$, де $\xi_n = \xi(t_n)$.

Так само, як і окрема випадкова величина, випадковий процес характеризується з максимально можливою повнотою своєю функцією розподілу ймовірностей (для простоти запису нехай n буде скінченим)

$$\begin{aligned} F_{\xi_1, \xi_2, \dots, \xi_n}(x(t_1); x(t_2); \dots; x(t_n)) &= \\ &= F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = \\ &= P(\xi_1 < x(t_1), \xi_2 < x(t_2), \dots, \xi_n < x(t_n)) = \\ &= P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n), \\ & \quad x_n = x(t_n). \end{aligned}$$

Тут у правій частині рівняння виписані ймовірності того, що в момент часу t_1 значення випадкової величини ξ_1 є меншим за x_1 , в момент часу t_2 значення випадкової величини ξ_2 є меншим за x_2 , в момент часу t_n значення випадкової величини ξ_n є меншим за x_n . Поза тим, $t_1 \leq t_2 \leq \dots \leq t_n$. Таким чином, випадковий процес описується функцією розподілу зліченої кількості випадкових величин

$(n \rightarrow \infty)$, або відповідною густиною розподілу ймовірностей. Зв'язок між цими характеристиками випадкового процесу наступний

$$F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = \\ = \int_{-\infty}^{x_1} dy_1 \int_{-\infty}^{x_2} dy_2 \dots \int_{-\infty}^{x_n} dy_n P_{\xi_1, \xi_2, \dots, \xi_n}(y_1, y_2, \dots, y_n).$$

Багатовимірна функція розподілу має наступний сенс

$$dP(x_1, x_2, \dots, x_n) = \\ F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n,$$

тобто це є ймовірність того, що в момент часу t_1 випадковий процес (випадкова функція) набуває значень з інтервалу $\xi_1 \in [x_1, x_1 + dx_1]$, в момент часу t_2 випадковий процес (випадкова функція) набуває значень з інтервалу $\xi_2 \in [x_2, x_2 + dx_2]$, ..., в момент часу t_n випадковий процес (випадкова функція) набуває значень з інтервалу $\xi_n \in [x_n, x_n + dx_n]$. Багатовимірна густина ймовірності задовольняє умові узгодженості

$$\int_{-\infty}^{\infty} dx_n P_{\xi_1, \xi_2, \dots, \xi_{n-1}, \xi_n}(x_1, x_2, \dots, x_{n-1}, x_n) = P_{\xi_1, \xi_2, \dots, \xi_{n-1}}(x_1, x_2, \dots, x_{n-1}).$$

Такій же умові задовольняє і багатовимірна функція розподілу ймовірностей

$$F_{\xi_1, \xi_2, \dots, \xi_{n-1}, \xi_n}(x_1, x_2, \dots, x_{n-1}, \infty) = F_{\xi_1, \xi_2, \dots, \xi_{n-1}}(x_1, x_2, \dots, x_{n-1}),$$

зокрема, ми отримаємо одновимірну (одноточкову) функцію розподілу ймовірностей у разі

$$F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, \infty, \dots, \infty) = F_{\xi_1}(x_1).$$

Аналогічно

$$\int_{-\infty}^{\infty} dx_2 \dots \int_{-\infty}^{\infty} dx_n P_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = P_{\xi_1}(x_1).$$

За означенням одноточкової густини розподілу має місце умова нормування

$$\int_{-\infty}^{\infty} dx P_{\xi}(x) = 1.$$

Умова нормування має місце і для густини ймовірностей довільної розмірності

$$\int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \dots \int_{-\infty}^{\infty} dx_n P_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = 1.$$

Загалом, чим меншою є розмірність функції розподілу ймовірностей, тим меншу інформацію, взагалі кажучи, вона містить про випадковий процес.

У разі збіжності деяких часових аргументів густина розподілу ймовірностей має стандартний щодо цих часів вигляд, наприклад,

$$P_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = \delta(x_1 - x_2) P_{\xi_2, \xi_3, \dots, \xi_n}(x_2, x_3, \dots, x_n).$$

Тут $t_1 \rightarrow t_2$. Нагадаємо, що $x_1 = x(t_1)$, $x_2 = x(t_2)$. Аналогічним чином можна взяти довільну пару аргументів вихідної густини розподілу, так само як і довільну кількість пар.

Умову нормування можна записати і для багатоточкової функції розподілу ймовірностей

$$F_{\xi_1, \xi_2, \dots, \xi_n}(\infty, \infty, \dots, \infty) = 1.$$

У багатьох сенсах густина розподілу ймовірностей є зручнішою характеристикою випадкового процесу, ніж функція розподілу ймовірностей. Саме густину зручно використовувати для знаходження числових характеристик випадкового процесу: математичного очікування, дисперсії тощо. Саме для густини розподілу ймовірностей можна природним чином здійснити її подальше узагальнення, ввівши умовну густину розподілу ймовірностей. За означенням умовної ймовірності

$$\begin{aligned} & P_{\xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k} | \xi_1, \xi_2, \dots, \xi_l}(x_{l+1}, x_{l+2}, \dots, x_{l+k} | x_1, x_2, \dots, x_l) = \\ &= \frac{P_{\xi_1, \xi_2, \dots, \xi_l, \xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k}}(x_1, x_2, \dots, x_l, x_{l+1}, x_{l+2}, \dots, x_{l+k})}{P_{\xi_1, \xi_2, \dots, \xi_l}(x_1, x_2, \dots, x_l)}. \end{aligned}$$

Тобто це є ймовірність того, що випадковий процес (випадкова функція) в момент часу t_{l+1} набуває значень з інтервалу $\xi_{l+1} \in [x_{l+1}, x_{l+1} + dx_{l+1}]$, в момент часу t_{l+2} набуває значень з інтервалу $\xi_{l+2} \in [x_{l+2}, x_{l+2} + dx_{l+2}]$, ..., в момент часу t_{l+k} випадковий процес (випадкова функція) набуває значень з інтервалу $\xi_{l+k} \in [x_{l+k}, x_{l+k} + dx_{l+k}]$ при умові, що в момент часу t_l він набуває значень з інтервалу $\xi_l \in [x_l, x_l + dx_l]$, в момент часу t_2 набуває значень з інтервалу $\xi_2 \in [x_2, x_2 + dx_2]$, ..., в момент часу t_1 випадковий процес (випадкова функція) набуває значень з інтервалу $\xi_l \in [x_l, x_l + dx_l]$

$$\begin{aligned} & dP(x_{l+1}, x_{l+2}, \dots, x_{l+k} | x_1, x_2, \dots, x_l) = \\ &= P_{\xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k} | \xi_1, \xi_2, \dots, \xi_l}(x_{l+1}, x_{l+2}, \dots, x_{l+k} | x_1, x_2, \dots, x_l) dx_{l+1} dx_{l+2} \dots dx_{l+k}. \end{aligned}$$

Легко отримати умову узгодженості для умовної густини розподілу ймовірностей. Дійсно,

$$\int_{-\infty}^{\infty} P_{\xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k} | \xi_1, \xi_2, \dots, \xi_l}(x_{l+1}, x_{l+2}, \dots, x_{l+k} | x_1, x_2, \dots, x_l) dx_{l+k} =$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{P_{\xi_1, \xi_2, \dots, \xi_l, \xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k}}(x_1, x_2, \dots, x_l, x_{l+1}, x_{l+2}, \dots, x_{l+k})}{P_{\xi_1, \xi_2, \dots, \xi_l}(x_1, x_2, \dots, x_l)} dx_{l+k} = \\
&= \frac{P_{\xi_1, \xi_2, \dots, \xi_l, \xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k-1}}(x_1, x_2, \dots, x_l, x_{l+1}, x_{l+2}, \dots, x_{l+k-1})}{P_{\xi_1, \xi_2, \dots, \xi_l}(x_1, x_2, \dots, x_l)} = \\
&= P_{\xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k-1} | \xi_1, \xi_2, \dots, \xi_l}(x_{l+1}, x_{l+2}, \dots, x_{l+k-1} | x_1, x_2, \dots, x_l).
\end{aligned}$$

Це дає можливість отримати умову нормування для умовної густини розподілу ймовірностей

$$\begin{aligned}
&\int_{-\infty}^{\infty} dx_{l+1} \dots \int_{-\infty}^{\infty} dx_{l+k} P_{\xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k} | \xi_1, \xi_2, \dots, \xi_l}(x_{l+1}, x_{l+2}, \dots, x_{l+k} | x_1, x_2, \dots, x_l) = \\
&= \frac{1}{P_{\xi_1, \xi_2, \dots, \xi_l}(x_1, x_2, \dots, x_l)} \int_{-\infty}^{\infty} dx_{l+1} \dots \int_{-\infty}^{\infty} dx_{l+k} \times \\
&\times P_{\xi_1, \xi_2, \dots, \xi_l, \xi_{l+1}, \xi_{l+2}, \dots, \xi_{l+k}}(x_1, x_2, \dots, x_l, x_{l+1}, x_{l+2}, \dots, x_{l+k}) = \\
&= \frac{1}{P_{\xi_1, \xi_2, \dots, \xi_l}(x_1, x_2, \dots, x_l)} P_{\xi_1, \xi_2, \dots, \xi_l}(x_1, x_2, \dots, x_l) = 1.
\end{aligned}$$

У разі двоточнової умовної густини розподілу ймовірностей умова нормування має вигляд

$$\int_{-\infty}^{\infty} dx_2 P_{\xi_2 | \xi_1}(x_2 | x_1) = 1.$$

Розподіл Пуассона

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 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 $P_2(dt) = \dots = P_r(dt) = 0$.

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

$$\frac{dP_2(t)}{dt} + \mu P_2(t) = \mu^2 t \exp(-\mu t) .$$

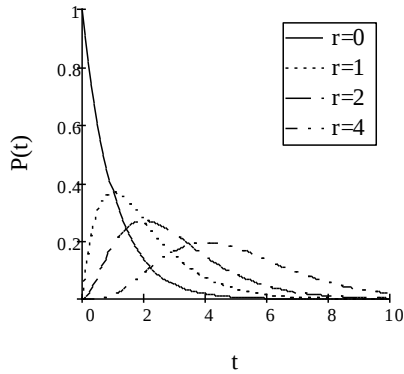
Розв'язавши його аналогічно попередньому, матимемо

$$P_2(t) = \frac{\mu^2 t^2}{2!} \exp(-\mu t) .$$

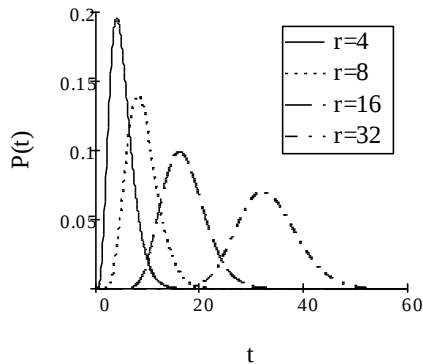
Продовжуючи цей процес отримання і розв'язання рівнянь до розгляду за проміжок часу $[0, t + dt]$ r подій, отримаємо

$$P_r(t) = \frac{\mu^r t^r}{r!} \exp(-\mu t) .$$

Наступний графік показує траєкторії розподілу Пуассона для різних значень випадкової величини



Малюнок 1. Траєкторії випадкового процесу для малих значень випадкової величини.



Малюнок 2. Траєкторії випадкового процесу для великих значень випадкової величини.

З малюнків видно, що траєкторії випадкового процесу Пуассона із зростанням випадкової величини дуже швидко наближаються до нормального розподілу. Коли $r = 32$ ці розподіли візуально не розрізняються.

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, 352 с.
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, 352 с.
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.

