С.В. Говаленков, В.М. Комяк,

Л.В. Мігунова, О.А.Тарасенко

Теорія ймовірностей
і математична статистика

Рекомендовано Міністерством освіти і науки України
як навчальний посібник для студентів вищих навчальних закладів

Харків 2003

УДК 51 (075.8)

Рецензенти: д. ф.-м. н., проф. Ольшанський В.П.
д. ф.-м. н., проф. Новожилова М.В.
д. т. н., проф. Путятін В.П.

Теорія ймовірностей і математична статистика. Навчальний посібник. / Говаленков С.В., Комяк В.М., Мігунова Л.В., Тарасенко О.А. - Харків: АПБУ, 2003. - 109 с.

Навчальний посібник розроблено для вивчення дисципліни “Теорія ймовірностей і математична статистика” в вищих навчальних закладах України пожежно-технічного профілю.
Учбовий матеріал розподілено на 3 розділи: випадкові події, випадкові величини, математична статистика. В першому розділі вивчаються основні поняття теорії множин, простір елементарних подій, класичне означення ймовірності, теореми додавання і множення, формула повної ймовірності. У другому розділі розглядаються випадкові величини та основні закони їх розподілу, закон великих чисел, початкові поняття випадкового процесу. У розділі математичної статистики описані вибірковий метод та методи статистичної перевірки гіпотез. Подання матеріала таке, що посібник є доступним для широкого кола студентів, які володіють основами математичного аналізу, лінійною алгеброю і теорією звичайних диференціальних рівнянь.

Редактор Т.О. Філіна
Коректор К.В. Хорошилова

Теорія ймовірностей і математична статистика.

Навчальний посібник.

Укладачі: Говаленков Сергій Валентинович
Комяк Валентина Михайлівна
Мігунова Лариса Василівна
Тарасенко Олександр Андрійович

Підп. до друк. 21.01.03 Формат 60х84 1/16
Друк – ризограф Умовн.-друк. арк. 7,0
Тираж 300 прим. Вид.№ 4/01 Зам.№

Дільниця оперативної поліграфії АПБ України
61023 м. Харків, вул. Чернишевського, 94

ВСТУП

Теорія ймовірностей - це математична наука, яка вивчає закономірності, що властиві випадковим явищам. Як усяка математична наука, вона має аксіоматичну побудову, з якої виводяться подальші результати. Основні поняття теорії ймовірностей не довільні, а в загальній формі відбивають певні сторони реальної дійсності. Завдяки цьому і висновки, які одержані в теорії ймовірностей, мають практичну цінність.

Основними поняттями теорії ймовірностей є поняття випадкової події і випадкової величини. В залежності від об’єкта вивчення теорію ймовірностей можна розділити на теорію ймовірностей випадкових подій (класичну теорію ймовірностей) і на теорію ймовірностей випадкових величин (сучасну теорію ймовірностей).

Випадковим називається всяке явище, характер протікання якого не можна цілком передбачити на підставі наявних у нас даних. Неможливість передбачення не означає відсутності причинного зв'язку між початковими даними і результатом. Вона викликана неповною поінформованістю про цей зв'язок. Проте неповнота знань не є перешкодою для з'ясування загальних закономірностей, які властиві випадковим явищам. Експериментатору добре відома наступна універсальна схема: чим більше випробувань, тим достовірніше може бути виведена закономірність і тим менша роль випадкових відхилень.

Теорія ймовірностей вивчає масові випадкові явища, тобто явища, що допускають хоча б у принципі експериментальну перевірку в однотипних умовах необмежене число разів. При цьому розглядаються такі випадкові явища, об'єктивні характеристики яких можуть бути отримані з будь-яким ступенем точності при будь-якому необмеженому повторенні експерименту.

1. ВИПАДКОВІ ПОДІЇ

1.1. Випробування. Простір елементарних подій. Випадкові події. Випробування. Простір елементарних подій

Під випробуванням у теорії ймовірностей розуміється експеримент, що може бути повторений необмежене число раз при дотриманні визначеного комплексу умов.

Завдання комплексу умов не вичерпує всіх обставин, що впливають на результат експерименту, тому при повторенні випробування може спостерігатися різний результат експерименту.

Наприклад, експеримент полягає у вийманні навмання кульки з урни, що містить M кульок, серед яких m білих кульок і M-m чорних. Комплекс умов: склад кульок за кольором; витаскування кульки навмання.

Експеримент може бути повторений нескінчену кількість разів, якщо вийняті кульки повертати назад. Даний експеримент можна назвати випробуванням.

Змінюючи комплекс заданих умов, що характеризують випробування, одержимо нове випробування.

Для кожного випробування можна вказати деяку систему можливих результатів, які мають ту властивість, що в результаті випробування повинний здійснитися один і тільки один результат. Така система можливих результатів, що пов'язана з даним випробуванням, називається простором елементарних подій Ω, а результати, що її складають - елементарними подіями ω.

Приклад 1. Випробування – вибірка кульки з урни, що містить m білих і M-m чорних кульок, з поверненням .

Можливі результати: ω1 – витягнена біла кулька,

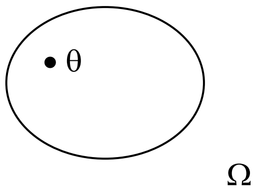
ω2 – витягнена чорна кулька,

Ω=(ω1, ω2) – кінцева множина.

Приклад 2. Випробування – стрільба по мішені до першого влучення Ω(ω1, ω2,…,ωn…); ωi–номер i-го влучення. ω∞ – влучення не відбувається.

Ω – нескінченна зчислена множина елементарних подій.

Приклад 3. Постріл по мішені з гарантованим влученням. Розміром кульки можна зневажити. Ставимо у відповідність кожній точці мішені θ результат випробування ω(θ). Ω містить нескінченну незчисленну множину елементарних подій.



Випадкова подія. Алгебра випадкових подій

Хай Ω – простір елементарних подій. Множина елементарних подій називається випадковою подією. Події позначають великими літерами латинського алфавиту: A,B,C,... Про елементарні події, що входять у випадкову подію, говорять, що вони їй сприяють.

Ω, ω – випадкові події.

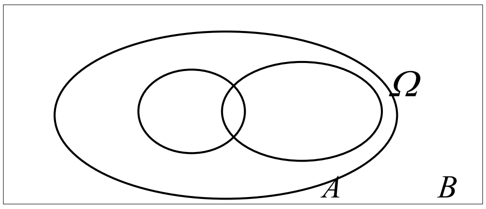
Ω обов'язково виникає в результаті випробування, тому Ω називається достовірною подією.

Приклад 4. В урні знаходяться дві білі і дві чорні кулі. Перенумеруємо: ω1, ω2 – білі кулі; ω3, ω4 – чорні кулі.

Ω = (ω1, ω2, ω3, ω4). A(ω1, ω3) – випадкова подія, що полягає в виборі кулі з непарним номером.

B (ω2, ω1) – подія – вибір білої кулі.

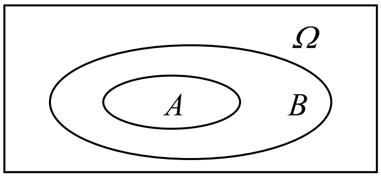
Геометрично випадкові події зображуються множиною точок Ω.



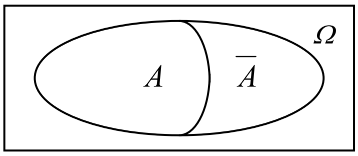
Алгебра випадкових подій (у рамках простору Ω)

1. Дві події називаються рівносильними (тотожними), якщо вони складаються з одних і тих же елементарних подій (A=B).

2. Подія B називається наслідком події A, якщо з появи події A випливає поява B. $A\subset B$.
Якщо $A\subset B$; $B\subset A$, то A=B; $A\subset B$; $B\subset C$; $A\subset C$.



3. Протилежною подією $\overline{A}$ називається подія, що рівносильна непояві A.
$\overline{\overline{A}}$ =A; $A\subset B$; $\overline{B}\subset\overline{A}$.



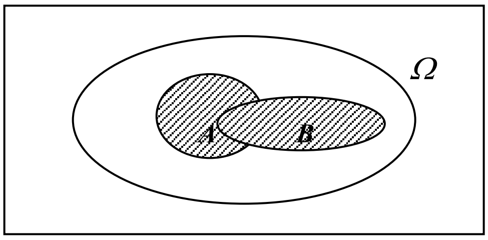
Приклад 4. Відбувся виклик пожежного відділення на пожежу. Подія A – виклик дійсний, подія $\overline{A}$ - виклик хибний.

$\overline{\Omega}$ – неможлива подія – "порожня множина $\varnothing$", $\overline{\Omega}\subset$A.

4. Сумою двох подій A і B називається подія A + B чи $A\cup B$, що означає, що в результаті випробування відбудеться принаймні одна з подій A чи B.

Зокрема, $A\cup\overline{A}=\Omega$, $A\cup\varnothing$=A

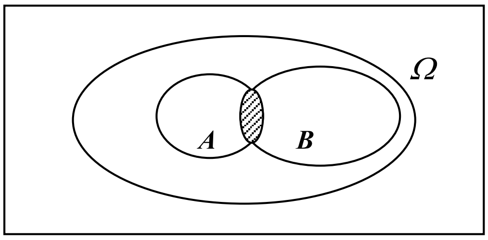
5. Добутком (сполученням) A·B чи $A\cap B$ називається подія, що складається з елементарних подій, які сприяють і A, і B, тобто одночасно відбуваються дві події.



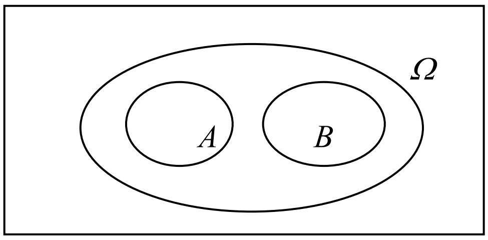
Зокрема, $A\cap\varnothing=\varnothing$, $A\cap\overline{A}=\varnothing$

Поняття суми і добутку подій поширюється на будь-яке число подій, як скінченне, так і нескінченне.

6. Події А та В називаються несумісними, якщо вони не можуть з'явитися в одному випробуванні. $AB=\overline{\Omega}$.

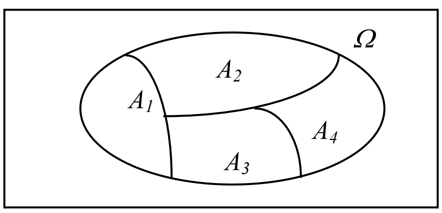


7. Події називаються попарно несумісними, якщо будь-які дві з них - несумісні.



Події А1, А2,, …, Аn, складають повну групу, якщо вони попарно несумісні, а сума їх дає достовірну подію.

Геометрично область Ω поділяється на області А1, А2,, …, Аn, що не мають попарно загальних точок.



Основні формули алгебри випадкових подій

Комутативність додавання і множення: A+B=B+A; AB=BA.

Асоціативність додавання і множення: A+(B+C)=(A+B)+C; A(BC)=(AB)C.

$AB\subset A\subset A+B$.

Якщо відбувається подія A, то відбувається і подія B.

$A\subset B\Rightarrow A+B=B$, $AB=A$; $A+\Omega=\Omega$; $A\Omega=A$; $A+\overline{\Omega}=A$; $A\overline{\Omega}=\overline{\Omega}$; A+A=A; AA=A.

Дистрибутивність додавання щодо множення:

(A+B)C=AC+BC;

дистрибутивність множення щодо додавання :

(AB)+C=(A+C)(B+C)

$$B\left(\sum_{k=1}^{n} A_k\right)=\sum_{k=1}^{n} BA_k$$

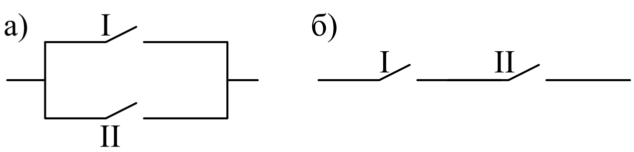
Закон подвійності. При переході до протилежних подій сума замінюється добутком і навпаки:

$$\overline{A+B}=\overline{A}\cdot\overline{B}, \qquad \overline{\sum_{k=1}^{n} A_k}=\prod_{k=1}^{n}\overline{A}_k,$$

$$\overline{AB}=\overline{A}+\overline{B}., \overline{A_1\cdot A_2\ldots A_n}=\sum_{k=1}^{n}\overline{A_k};$$

;

Приклад 5. Електричний ланцюг, що містить два контакти, які з'єднані послідовно і паралельно



$\Omega=(\omega1,\omega2,\omega3,\omega4)$:

ω1– обидва замкнені;

ω2 – обидва розімкнені;

ω3 – I- замкнений; II- розімкнений;

ω4 – I-розімкнений; II- замкнений.

Події: A - I контакт замкнений; B - II контакт замкнений; C - ланцюг замкнений.

$A=\omega1+\omega3$; $B=\omega1+\omega4$;

а) C=A+B; $\overline{C}=\overline{A}\,\overline{B};$ b) C=AB; $\overline{C}=\overline{A}+\overline{B}.$

ЗАДАЧІ

1.1.1. Нехай A, B, C – три довільні події, які полягають у тому, що відбулись пожежі I номеру виклику, II номеру виклику, III номеру виклику відповідно. Знайти вираз для подій, що

а) відбулася тільки пожежа I номеру виклику;

б) відбулися пожежі тільки I і II номерів виклику;

в) відбулися пожежі усіх трьох номерів виклику ;

г) відбулася принаймні одна з пожеж;

д) відбулися пожежі принаймні двох номерів виклику;

е) відбулася пожежа тілки одного номеру виклику;

ж) відбулися пожежі двох і тільки двох номерів виклику;

з) жодна з пожеж не відбулася;

і) відбулось пожеж не більше двох номерів виклику.

Відповідь:

а) $A\overline{B}\,\overline{C}$; *б)* $AB\overline{C}$; *в)* ABC; *г)* $\Omega - \overline{A}\,\overline{B}\,\overline{C}$; *д)* $\Omega - \overline{A}\,\overline{B}\,\overline{C} - A\overline{B}\,\overline{C} - \overline{A}\,\overline{B}C - \overline{A}B\overline{C}$; *е)* $A\overline{B}\,\overline{C} + B\overline{A}\,\overline{C} + C\overline{A}\,\overline{B}$; *ж)* $AB\overline{C} + AC\overline{B} + BC\overline{A}$; *з)* $\overline{A}\,\overline{B}\,\overline{C}$; *і)* $\Omega - ABC$.

1.1.2. З множини подружніх пар навмання вибирається одна пара. Подія A: “чоловіку більше 30 років”, подія B: “чоловік старший за дружину”, подія C: “дружині більше 30 років”.

а) З'ясувати зміст подій ABC, A$\overline{B}$, $A\overline{B}C$.

б) Перевірити, що $A\overline{C} \subset B$.

Відповідь: а) ABC – “обом більше 30 років, причому чоловік старший за дружину”; A-AB – “чоловікові більше 30 років, але він не старший за свою дружину”; $A\overline{B}C$ – “обом більше 30 років, причому чоловік не старший за свою дружину”;

б) $A\overline{C}$ – “чоловіку більше 30 років” і “дружині не більше 30 років”, отже, чоловік старший за дружину – B, тобто $A\overline{C} \subset B$.

1.1.3. Яка умова сумісності подій A+B, $\overline{A} + B$ і $A + \overline{B}$?

Відповідь: Оскільки (A+B)($\overline{A} + B)(A + \overline{B}) = AB$, то потрібна сумісність A і B.

1.1.4. Довести, що подія (A+B)$(A + \overline{B})(\overline{A} + B)(\overline{A} + \overline{B})$ неможлива.

1.1.5. Чи рівносильні події A і B, якщо:

а) $\overline{A} = \overline{B}$?

б) A+C=B+C?

в) AC=BC?

Відповідь: а) так; б), в) – у загальному випадку - ні.

1.1.6. Нехай А, В, С – довільні події. Спростити наступні вирази для подій:

а) (А+В) (В+С);

б) (А+В) (А+$\overline{B}$).

Відповідь: а) (А+В) (В+С) =АВ+АС+ВВ+ВС=(А+В+С)В+АС=В+АС;

б) (А+В)(А+$\overline{B}$)= АА+АВ+А$\overline{B}$ +В$\overline{B}$ =А+А(В+$\overline{B}$)+$\overline{\Omega}$ = А+АΩ+$\overline{\Omega}$ =А.

1.2 Імовірність. Аксіоми теорії ймовірностей. Теореми додавання і множення. Умовна ймовірність Імовірність. Аксіоми теорії ймовірностей

Імовірність – числова міра ступеня об'єктивної можливості випадкової події А. Розглянемо аксіоми теорії ймовірностей у припущенні, що простір елементарних подій дискретний, тобто складається із скінченного або счисленного числа елементарних подій ω1, ω2,…,ωn…

Аксіома 1

З кожною елементарною подією ω_i зв'язане невід'ємне число $p(\omega_i)$, яке називається ймовірністю цієї елементарної події.

Аксіома 2

Ці числа такі, що їх сума дорівнює одиниці, тобто

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

Аксіома 3

Імовірність $p(A)$ будь якої випадкової події A дорівнює сумі ймовірностей елементарних подій, із яких складається подія

$$A:\ p(A)=\sum_{\omega_I \in A} P(\omega_I).$$

Із цих аксіом витікають такі висновки:

Висновок 1

Імовірність достовірної події дорівнює одиниці (умова нормування):

p (Ω) =1.

Висновок 2

Імовірність неможливої події дорівнює нулю:

$p(\overline{\Omega})=0.$

Висновок 3

Імовірність будь-якої випадкової події A міститься між нулем і одиницею:

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

Висновок 4

Правило додавання:

Якщо A1, A2,…, An – попарно несумісні події, то

p(A1+ A2+…+An)=p(A1) + p(A2)+…+p(An).

Висновок 4’

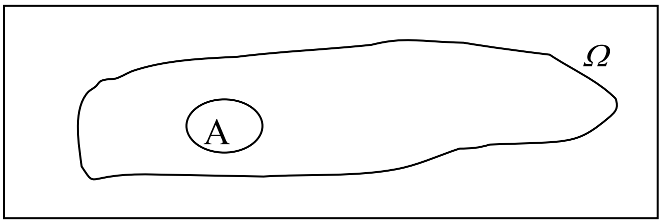
Якщо випадкові події A1, A2,…, An…– попарно несумісні події, то

$$p\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} p(A_k).$$

Висновок 5

Нехай $A \subset B$, то $p(A) \le p(B)$.

Оскільки $B=A+\overline{A}\cdot B$, то $p(B)=p(A)+p(\overline{A}\cdot B) \ge p(A)$, звідси $p(B)>p(A)$.



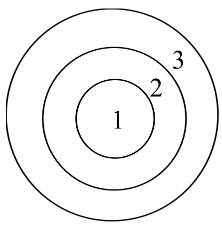
Висновок 6

Якщо випадкові події A1, A2,…, An утворюють повну групу подій, то:

$$p\left(\sum_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} p(A_k) = 1;$$

$\Omega=\sum_{k=1}^{n} A_k;$ $1 = p(\Omega)=\sum_{k=1}^{n} p(A_k);$ $p(\overline{A})=1-p(A);$ $\Omega=A+\overline{A};$ $1=p(A)+p(\overline{A})$.

Приклад 1. Кругла мішень має три зони: 1, 2, 3. Імовірність попадання в першу зону при одному пострілі дорівнює 0,15; в другу – 0,23; в третю – 0,37. Знайти ймовірність промаху.



Позначимо через А – промах, Ā – попадання, А1-попадання в першу зону, А2 – попадання в другу зону, А3 -.попадання в третю зону Тоді A=A1+A2+A3.

P(Ā)=P(A1)+P(A2)+P(A3)=0,15+0,23+0,37=0,75

Події A і Ā створюють повну групу, тобто A+Ā=Ω. Згідно з висновком 6:

P(A)=1-P(Ā)=1-0,75=0,25

Висновок 7

У випадку сумісних подій

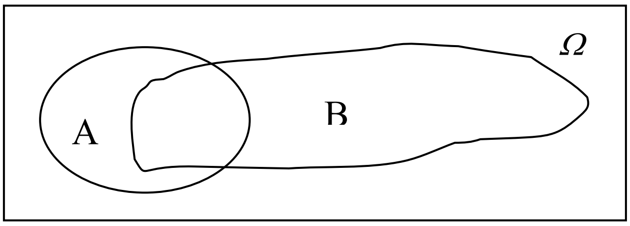
p(A+B)=p(A)+p(B)-p(AB).

Оскільки A+B=AB+$\bar{A}B$ +A$\bar{B}$, то $p(A+B)=p(AB)+p(\bar{A}B)+p(A\bar{B});$

A=AB+A$\bar{B}$, p(A)=p(AB)+p(A$\bar{B}$);

B=AB+$\bar{A}B$,p(B)=p(AB)+p($\bar{A}B)$;

p(A+B)=p(A)+p(B)-p(AB).



У випадку несумісних подій AB=$\overline{\Omega}$; p(AB)=0.

Способи підрахунку ймовірностей

1. Схема урн (класична схема)

Хай Ω складається з скінченного числа рівноймовірних елементарних подій.

Ω(ω1, ω2,…,ωn); p(ω1)=p(ω2)=…=p(ωn)= p(ω); Ω=ω1+ ω2+…+ωn;

1=p(ω1)+p(ω2)…+p(ωn)= np(ω); $p(\omega_i)=\frac{1}{n};\ i=\overline{1,n}.$

Нехай A=ω1+ ω2+…+ωm, тоді

p(A)= p(ω1)+p(ω2)…+p(ωm) = $\frac{m}{n}$.

Класичне визначення ймовірності:

Імовірність випадкової події А дорівнює відношенню числа m сприятливих їй елементарних подій до їхнього загального числа n:

$$p(A)=\frac{m}{n}.$$

2. Статистична ймовірність

Частість – це кількісна характеристика випадкових подій.

Нехай у серії з n випробувань m раз з'являється подія А : 0≤m≤ n.

Відношення $\frac{m}{n}$ називається частістю появи події А в даній серії випробувань і позначається

$$p^*(A)=\frac{m}{n}.$$

Частість має наступні властивості:

2.1. Для всякої події А і всякої серії випробувань

0≤p*(A)≤1.

2.2. $p^*(\Omega)=1.$

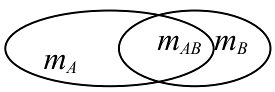
2.3. Якщо події А і В несумісні, то

$$p^*(A+B)=p^*(A)+p^*(B).$$

Нехай при n випробуваннях отримано m_A появ події А и m_B появ події В, тоді m_A+m_B – число появи події А+В.

$$p^*(A+B)=\frac{m_A+m_B}{n}=\frac{m_A}{n}+\frac{m_B}{n}=p^*(A)+p^*(B).$$

Якщо А і В – сумісні, то



$$p^*(A+B)<p^*(A)+p^*(B).$$

$$p^*(A+B)=\frac{m_A+m_B-m_{AB}}{n}<p^*(A)+p^*(B).$$

Поняття частість є основним при експериментальному вивченні випадкових подій. Однак частота не може служити об'єктивною характеристикою випадкової події, яка досліджується, тому що залежить від випадкового збігу обставин, що пов'язані з даною серією випробувань, та від індивідуальних особливостей самого експериментатора. Однак зі збільшенням числа випробувань частість стає “сталою”.

Імовірність випадкової події відповідає в ідеалізованому виді тій межі, до якої має тенденцію прагнути стала частота події при необмеженому збільшенні числа випробувань. Ця межа носить назву статистичної ймовірності.

Теорія ймовірностей призначена для опису випадкових подій, що мають сталу частість.

3. Геометрична ймовірність

Якщо Ω(ω) містить незчисленну множину елементарних подій, то схема урн не застосується. Однак у деяких задачах удається Ω(ω) розглянути як частину звичайного геометричного простору, а ω – випадкове влучення точки простору в область Ω(ω).

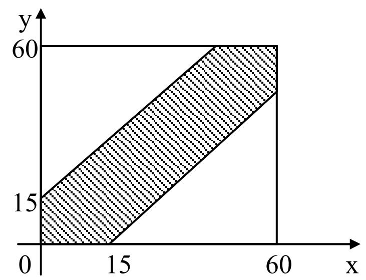
Теорема. Нехай ймовірність P(A) влучення в будь-яку область A, що лежить усередині Ω, пропорційна площі цієї області і не залежить від форми A і розташування A усередині Ω. Тоді ймовірність P(A) дорівнює відношенню площі області, зайнятої сприятливими положеннями точки влучення, до площі області Ω

$P(A)=\dfrac{S_A}{S_\Omega}$,

P(A)=kSA, P(Ω)=kSΩ=1; k= $\dfrac{1}{S_\Omega}$, P(A)= $\dfrac{S_A}{S_\Omega}$.

Зауваження. P(M)=0 (ймовірність точки), проте не випливає, що воно неможливо.

Приклад 2. Дві особи домовляються про зустріч, що повинна відбутися між 12 і 13 годинами у визначеному місці. Перший, що прийшов, очікує іншого протягом τ хвилин, а потім іде. Момент приходу кожної особи з рівним успіхом може виявитися будь-яким в інтервалі (0, 60). Знайти ймовірність їхньої зустрічі. -



Випробування можна розглядати як випадкове кидання точки (x, y) у квадрат.

Сприятлива область A: |x-y| ≤ τ.

$$P(A) = \frac{60^2 - (60-\tau)^2}{60^2};$$

$$\tau = 15; \qquad P(A) = \frac{60^2 - 45^2}{60^2} = \frac{7}{16}.$$

4. Комбінаторні методи

Прямий підрахунок ймовірностей іноді виявляється громіздким. У цьому випадку використовують комбінаторні методи. Комбінаторика – теорія скінченних множин.

Задано дві множини елементів:

a1, a2, …am1;

b1, b2, ,,,bm2;

	a1	…	am1
b1	b1a1	…	b1am1
…	…	…	…
bm2	bm2a1	…	bm2am1

4.1. Число способів вибору з кожної множини по одному елементу – m1m2 (основне правило комбінаторики).

4.2. Задано k множин елементів (груп) по одному елементу з кожної групи m1m2…mk...

4.3. Число способів вибору k елементів із множини a1, a2,…,am по одному з поверненням, якщо вважати різними підмножини елементів, що відрізняються порядком елементів, дорівнює $m \cdot m \cdot \ldots \cdot m = m^k$.

Приклад 3. Розглянемо множину елементів $A\{a,b,c\}$. Скількома способами можна вибрати з трьох елементів $(m=3)$ підмножини із 2 елементів $(k=2)$ по схемі з поверненням. Позначимо число таких способів через $M_2(k) = 3^2 = 9$. Отримані підмножини називають розміщення з повторенням і вони мають вигляд: $\{a,b\},\{b,c\},\{a,c\},\{b,a\},\{c,b\},\{c,a\},\{a,a\},\{b,b\},\{c,c\}$

4.4. Якщо підмножини утворюються по схемі без повернення елементів, то вони утворюються $m\cdot(m-1)\cdot\ldots\cdot(m-k+1)$ способами, або $A_m^k = \frac{m!}{(m-k)!}$. Для приклада 1 число таких способів дорівнює $M_2(k) = 3\cdot 2 = 6$. Отримані підмножини називають розміщеннями із 3 елементів по 2 і вони мають вигляд $\{a,b\},\{b,c\},\{a,c\},\{b,a\},\{c,b\},\{c,a\}$. Це підмножини, які враховують порядок розташування елементів.

4.5. Розміщення із m елементів по m називають перестановками і позначають P_m. Кількість перестановок із m елементів обчислюється за формулою $m!$ Перестановки обчислюють упорядковані множини, тобто множини, які відрізняються тільки порядком їх елементів. Розглянемо перестановки множини А:(a, b, c),(a, c, b),(b, a, c),(b, c, a), (c, b, a),(c, a, b).

4.6. Часто розглядають підмножини елементів об'ємом k елементів із множини a1, a2,…,am (вибір без повернення елементів), які відрізняються між собою хоча б одним елементом (порядок елементів не враховується). Такі підмножини носять назву комбінацій із m елементів по k і позначаються C_m^k. Кількість комбінацій із m елементів по k знаходиться за формулою

$$C_m^k = \frac{m!}{k!(m-k)!}.$$

Наприклад, із множин $\{a,b,c\}$ утворити двух (к=2) елементні підмножини, тобто $C_3^2 = 3$. Це підмножини: $\{a,b\},\{a,c\},\{b,c\}$.

4.7. Число розбивок множини з n елементів на k підмножин (перестановки з повтореннями), що містять m1, m2,…mk елементів (m1+m2+…+mk=n), дорівнює

$$C_n(m_1,m_2,...,m_k) = \frac{n!}{m_1!...m_k!}.$$

Приклад4: визначити число різних слів, переставляючи літери в слові "математика".

n=10 – кількість літер в слові "математика", m1=2 – літера М, m2=3 – літера А,m3=2 – літера Т,m4=1 (Е),m5=1 (И),m6=1 (К)

C10(2, 3, 2, 1, 1, 1)= $\frac{10!}{2!\cdot 3!\cdot 2!\cdot 1!\cdot 1!\cdot 1!} = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9}{2\cdot 2\cdot 3\cdot 2} = 151200$

4.8. Часто розглядають підмножини елементів об'ємом k із множини a1, a2,…,am (вибір з повернення елементів), які називають комбінаціями з повторенням.

Комбінації із n елементів по m елементів називають всі підмножини з m елементів без врахування порядку їх елементів.

Число різних комбінацій із n елементів по m з повторенням дорівнює

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

Наприклад, з трьох елементів (a, b, c) можна скласти такі комбінації по 2: (a, a), (a, c), (b, c), (a, b), (b, b), (c, c).

Приклад 5. В картотеці зі звітами про виклики на пожежі М карток, m - про дійсні виклики, (M – m) - про хибні. З картотеки по черзі з поверненням виймають дві картки. Яка ймовірність, що обидві картки про дійсні виклики ?

Події: А – вибір 1-ї картки про дійсний виклик,

В – вибір 2-ї картки про дійсний виклик.

Елементарні події повинні бути рівноймовірними.

Усіх вибірок M2, сприятливих - m2.

$$P(AB)=\frac{m^2}{M^2}.$$

Приклад 6. З урни, що містить М куль, з яких m білих, вибирають 2 кулі без повернення. Знайти ймовірність вибору обох білих куль.

За умовою порядок вибору куль не цікавить.

$$P(A)=\frac{C_m^2}{C_M^2}=\frac{m(m-1)}{M(M-1)}.$$

Умовна ймовірність. Правило множення ймовірностей

Нехай у серії з n випробувань mA разів відбулася подія А; mB разів– подія B; mAB разів– подія AB:

$$p^*(A)=\frac{m_A}{n}; \quad p^*(B)=\frac{m_B}{n}; \quad p^*(AB)=\frac{m_{AB}}{m_B}; \quad p^*(A/B)=\frac{p^*(AB)}{p^*(B)}.$$

Визначення 1. Умовною ймовірністю p(A/B) появи події А, за умови появи події В, називається відношення p(AB) сумісних подій до ймовірності p(B):

$$p(A/B)=\frac{p(AB)}{p(B)}, \qquad (1.2.1)$$

$$p(B/A)=\frac{p(AB)}{p(A)}. \qquad (1.2..2)$$

З (1.2.1) і (1.2.2) випливає правило множення ймовірностей:

p(AB)=p(A)p(B/A)=p(B)p(A/B), (1.2.3)

p(ABC)=p(A/BC)p(BC)=p(A/BC)p(B/C)p(C).

Зауваження: зміст визначення 1.

Вимога появи події В змінює комплекс умов і приводить до нового випробування. При цьому Ω/ = B; A/=AB.

Завдання полягає в такому перерахуванні p(A/), в результаті якого Ω/ задовольняє виразам (1.2.1) –(1.2.3).

Цього можна досягти зміною масштабу, змінюючи кожну ймовірність на p(B):

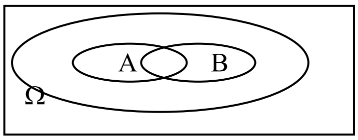
$$p(A/B)=\frac{p(AB)}{p(B)}.$$

Умовні ймовірності одержуються з початковим масштабуванням, отже, підкоряються усім властивостям теорії імовірностей.

Визначення 2. Випадкові події A і B називаються незалежними, якщо припущення про появу одної з них не впливає на ймовірність появи іншої.

p(A/B)=p(A), p(B/A)=p(B),

p(AB)=p(A)p(B) - для незалежних подій.



$$P(A)=\frac{S_A}{S_\Omega}$$

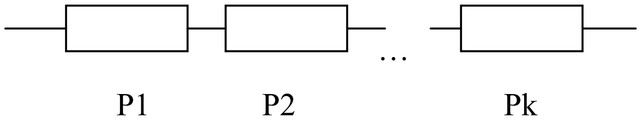
$$P(A/B)=\frac{S_{A'}}{S_{\Omega'}}=\frac{S_{A'}/S_\Omega}{S_{\Omega'}/S_\Omega}=\frac{P(AB)}{P(B)}$$

Приклад 7. Отримано партію з M вогнегасників. Навмання відбирають k вогнегасників (вибірка без повернення) і перевіряють їхню якість. Якщо серед відібраних вогнегасників не виявиться жодного бракованого – партію приймають, у противному випадку - повертають постачальнику. Знайти ймовірність того, що партія буде прийнята, якщо в ній міститься m бракованих вогнегасників.

Ai – подія, яка полягає в тому, що вогнегасник з номером і не бракований.

$$P(A_1..A_k)=P(A_1)P(A_2/A_1)P(A_3/A_1A_2)..P(A_k/A_1..A_{k-1})=\frac{M-m}{M}\cdot\frac{M-m-1}{M-1}\cdots\frac{M-m_1-(k-1)}{M-(k-1)}=\frac{C^k_{M-m}}{C^k_M}$$

Приклад 8. Знайти ймовірність виходу з ладу послідовно з'єднаного ланцюга, що містить k елементів; pi – ймовірність виходу з ладу і-го елемента. -



Ai – подія у виході і-го елемента (події незалежні).

B – вихід ланцюга.

B = A1+A2+…+Ak

$\overline{B} = \overline{A}_1 \overline{A}_2 ... \overline{A}_k$ – система працює.

$P(\overline{B}) = P(\overline{A}_1)...P(\overline{A}_k) = (1 - p_1)(1 - p_2)...(1 - p_k).$

$P(B) = 1 - (1 - p_1)(1 - p_2)...(1 - p_k)$.

Задачі

1.2.1. Командир пожежного відділення тричі викликає ЦУСЗ. Ймовірність того, що буде прийнятий перший виклик дорівнює 0,4, другий виклик – 0,3, третій виклик - 0,2. За умовами прийому події, які полягають в тому, що даний виклик буде почутий, незалежні. Знайти ймовірність того, що на ЦУСЗ взагалі почують виклик.

Відповідь: 0,664.

1.2.2. У змаганнях з пожежно-прикладного спорту беруть участь 20 чоловік, із яких 10 чоловік є найбільш сильними. По жеребу 20 чоловік розподіляються на 2 групи, відповідно по 8 і 12 чоловік. Знайти ймовірність того , що

а) двоє найбільш сильних гравців будуть змагатися в різних групах;

б) четверо найбільш сильних потраплять по двоє у різні групи.

Відповідь: а) $p_1 = \frac{C_8^1 \cdot C_{12}^1}{C_{20}^{10}} = \frac{24}{46189}$, б) $p_2 = \frac{C_8^2 \cdot C_{12}^2}{C_{20}^{10}} = \frac{7}{4199}$.

1.2.3. Томи чотиритомного твору розташовані на полиці у випадковому порядку. Знайти ймовірність того, що томи стоять у належному порядку справа наліво чи зліва направо.

Відповідь: р = $\frac{2}{24} = \frac{1}{12}$.

1.2.4. Для безперешкодного польоту над деякою територією літак, наближуючись до неї, посилає по радіо парольну кодову групу, що складається з декількох крапок і тире. Знайти ймовірність того, що радист, який не знає парольної групи, угадає її, передавши яку-небудь групу навмання, якщо відомо, що число кодових елементів у групі (крапок і тире): а) 5; б) 7.

Відповідь: а) р = $\frac{1}{32}$; б) р = $\frac{1}{128}$.

1.2.5. З колоди в 32 карти навмання вибираються 4. Знайти ймовірність того, що серед них виявиться хоча б один туз.

Відповідь: p=1-q (“4 не тузи”) = 1 - $\frac{C_{28}^4}{C_{32}^4} = 1 - \frac{4095}{7192} \approx 0,43$.

1.2.6. Нехай A і B – випадкові події. Доведіть тотожності:

а) $p(\overline{A}\,\overline{B}) = 1 - p(A) - p(B) + p(AB)$; б) $p(A) + p(\overline{A}B) = p(\overline{B}A)$.

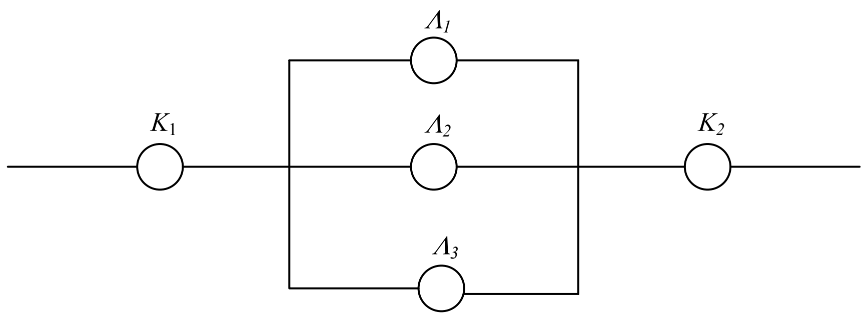
Відповідь: а) $p(\overline{A}\,\overline{B}) = p(\overline{A+B}) = 1 - p(A+B) = 1 - p(A) - p(B) + p(AB)$,

б) p(AB)=p(A)p(B/A)=p(B)p(A/B) чи

$p(AB) = p(A)(1 - p(\overline{B}/A) = p(B)(1 - p(\overline{A}/B)$,

звідки $p(A) + p(B\overline{A}) = p(B) + p(A\overline{B})$.

1.2.7. Електричний ланцюг MN складений за схемою



N

Вихід з ладу за час T різних елементів заданий таблицею

Елементи	K1	K2	Λ1	Λ2	Λ3
Ймовірність	0,6	0,5	0,4	0,7	0,9

Визначити ймовірності розриву ланцюга за час T.

Рішення. Імовірність розриву паралельного ланцюга Λ1, Λ2, Λ3 дорівнює

P(A)=0,4·0,7·0,9=0,252. Ймовірність розриву послідовно з'єднаного ланцюга, що складається з елементів K1, Λ, K2 дорівнює

P=1-(1-0,6)(1-0,252)(1-0,5)=1-0,4·0,748·0,5=0,1596.

1.2.8. У шухляді лежать 20 тенісних м'ячів, у тому числі 12 нових і 8 таких, з якими вже грали. Із шухляди витягаються навмання два м'ячі для гри і після гри повертаються в шухляду. Після цього із шухляди виймають два м'ячі для наступної гри. Знайти ймовірність того, що ці обидва м'ячі будуть не тими, з якими грали?

Відповідь: 0,279.

1.2.9. Три стрільці зробили залп, причому кулі вразили мішень. Знайти ймовірність того, що третій стрілець вразив мішень, якщо імовірності влучення в мішень першим, другим і третім стрільцями дорівнюють 0,6; 0,5 і 0,4, відповідно.

Рішення: Позначимо подію Ai – i -ий стрілець вразив мішень; i=1,2,3, а подія А – у результаті залпу два з трьох стрільців вразили мішень. Тоді

$$A = A_1 A_2 \overline{A}_3 + \overline{A}_1 A_2 A_3 + A_1 \overline{A}_2 A_3;$$
$$p(A) = p(A_1 A_2 \overline{A}_3) + p(\overline{A}_1 A_2 A_3) + p(A_1 \overline{A}_2 A_3) =$$
$$= 0{,}6 \cdot 0{,}5 \cdot 0{,}6 + 0{,}4 \cdot 0{,}5 \cdot 0{,}4 + 0{,}6 \cdot 0{,}5 \cdot 0{,}4 = 0{,}38.$$

Події А сприяє подія В: $B = \overline{A}_1 A_2 A_3 + A_1 \overline{A}_2 A_3 w$ (третій стрілець потрапив);

$$p(B) = p(\overline{A}_1 A_2 A_3) + p(A_1 \overline{A}_2 A_3) = 0{,}4 \cdot 0{,}5 \cdot 0{,}4 + 0{,}6 \cdot 0{,}5 \cdot 0{,}4 = 0{,}20.$$

Ймовірність того, що третій стрілець вразив мішень, дорівнює відношенню імовірносних мір подій В і А, тобто

$$p = \frac{p(B)}{p(A)} = \frac{0{,}20}{0{,}38} = \frac{10}{19}.$$

1.2.10. Дві з чотирьох незалежно працюючих ламп приладу відмовили. Знайти ймовірність того, що відмовили перша і друга лампи, якщо ймовірності відмови першої, другої, третьої і четвертої ламп дорівнюють p1=0,1; p2=0,2; p3=0,3 і p4=0,4, відповідно.

Відповідь: p=0,039.

1.3. Формули повної ймовірності і формули Байєса.

Формула повної ймовірності і формули Байєса

Нехай H0, H1…,Hn – повна група подій (система гіпотез),

А – деяка подія:

Ω = H0+H1+…Hn,

$$A\Omega = \sum_{i=0}^{n} AH_i.$$

Відомо, що подія А відбудеться за умовою появи однієї з гіпотез.

P(A)= $\sum_{i=0}^{n} P(AH_i) = \sum_{i=0}^{n} P(H_i) P(A/H_i),$

P(A)= $\sum_{i=0}^{n} P(H_i) P(A/H_i)$ - формула повної ймовірності.

Безумовні ймовірності $P(H_i), i=0,...,n$ - це дослідні або апріорні ймовірності.

Хай відомі апріорні ймовірності і хай подія A здійснилася, але невідомо з якою із гіпотез . Яка ймовірність, що це була подія Hi ?

$$P(H_i/A)=\frac{P(H_iA)}{P(A)}=\frac{P(A/H_i)P(H_i)}{P(A)},$$

$$P(H_i/A)=\frac{P(A/H_i)P(H_i)}{\sum_{i=0}^{n}P(A/H_i)P(H_i)}$$ – формули Байєса.

Формула Байєса дає числове значення нової ймовірності (післядослідної, апостеріорної) кожної з подій Hi після того, як з'явилася подія A (переоцінити ймовірності гіпотез після отримання нової інформації).

Контролем правильності обчислень може служити тотожність:

$$\sum_{i=0}^{n}P(H_i/A)=1.$$

Приклад 1. Перевіряються 5 пожежних сповіщувачів, про які відомо, що два з них працездатні, а три – ні. Спочатку без перевірки відкладається один сповіщувач. Яка імовірність, що наступним буде перівератися працездатний?

Подія A – наступний сповіщувач - працездатний.

Гіпотези: H1 – перший відкладений сповіщувач - працездатний;

H2 – перший відкладений сповіщувач - непрацездатний;

A=AH1+AH2

P(A)=P(H1)P(A/H1)+ P(H2)P(A/H2)= 2/5 $\frac{1}{4}+\frac{3}{5}\cdot\frac{2}{4}=\frac{2}{5}$.

Приклад 2. З урни, що містить дві білі і три чорних кулі, виймуть кулю, колір якої - невідомий. Наступна витягнута куля – біла. Яка імовірність, що перша - теж біла?

Подія A – друга витягнута куля - біла; H1, H2 – означають, що перша витягнута куля біла або чорна, відповідно.

$$P(H_1/A)=\frac{P(H_1)P(A/H_1)}{P(H_1)P(A/H_1)+P(H_2)P(A/H_2)}=\frac{2/5\cdot 1/4}{2/5}=1/4.$$

Задачі

1.3.1. З аналізу багаторічної статистики відомо, що відношення кількости пожеж третього номеру виклика до кількости пожеж четвертого номеру виклика $\frac{m}{n}$ становить , при цьому, у к % випадків пожежі третього номеру виклику, і 1 % випадків пожежі четвертого номеру виклика відбуваються з людськими жертвами. Яка імовірність, що наступна пожежа третього або четвертого номерів виклику видбудеться з жертвами ?

Відповідь: $\frac{m}{m+n}\frac{k}{100}+\frac{n}{m+n}\frac{l}{100}$

1.3.2. У спеціалізовану лікарню надходить у середньому 50% хворих із захворюванням K, 30% – із захворюваннями L, 20% – із захворюванням M. Ймовірність повного лікування хвороби K дорівнює 0,7; для хвороб L і M ці ймовірності, відповідно, дорівнюють 0,8 і 0,9. Хворий, що надійшов у лікарню, був виписаний здоровим. Знайти ймовірність того, що цей хворий страждав захворюванням K.

Відповідь: p = 5/11.

2 ВИПАДКОВІ ВЕЛИЧИНИ

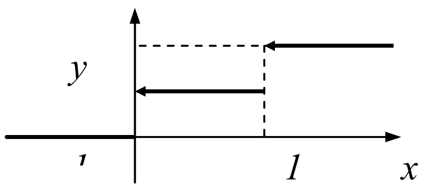
2.1 Випадкова величина. Функція розподілу випадкової величини

Поставимо у відповідність кожній елементарній події число. Наприклад, розглянемо випробування , що полягає в киданні монети. Можливі події: ω1 – випадання числа, ω2 – випадання герба. Хай f(ω1)=0, f(ω2)=1.

ξ= f(ω) – числова функція, що залежить від випадку.

Знаючи ймовірність P(ω1) і P(ω2) елементарних подій ω1 і ω2, легко обчислити для будь-якого значення x ймовірність P(ξ<x).

$$P(\xi<x)=\begin{cases}0, x\le 0,\\ P(\omega_1), 0<x\le 1,\\ 1, x>1.\end{cases}$$

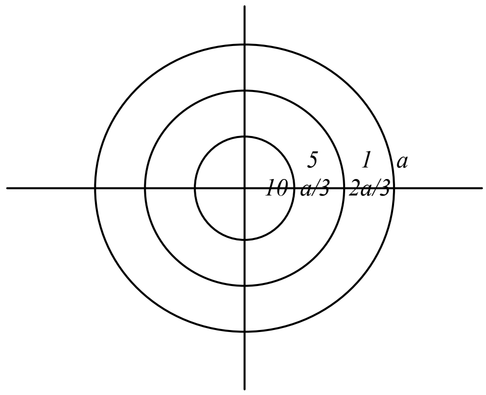


Нехай Ω – деякий простір елементарних подій. Випадковою величиною ξ називається всяка числова функція ξ=f(ω) елементарної події ω, для якої при будь-якому значенні x існує ймовірність нерівності ξ<x.

P(ξ<x), що розглянута як функція аргументу x, називається функцією розподілу випадкової величини ξ і позначається Fξ(x).

Випадкову величину будемо позначати грецькими буквами ξ,η, а значення, що вона набуває, латинськими x, y.

Наприклад: коло радіуса a – мішень з гарантованим влученням.



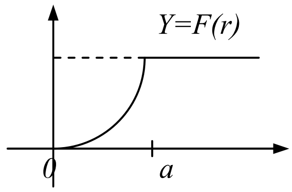
Випадкову величину можна ввести різними способами.

ξ – абсциса точки влучення ; $0 \le x \le a$;

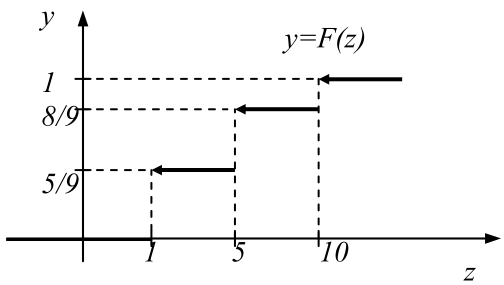
ρ- відстань від точки влучення до центра мішені; $0 \le \rho \le a$;

ζ- кількість очків, що відповідають точці влучення; $\zeta=1;5;10$.

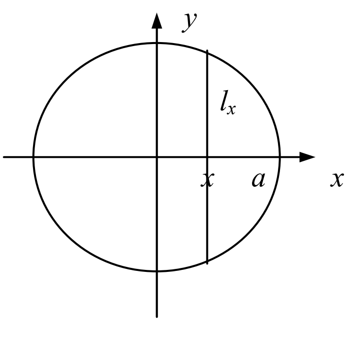
Fρ(r)=P(ρ<r)= $\begin{cases} 0, r \le 0, \\ \dfrac{\pi r^2}{\pi a^2} = \dfrac{r^2}{a^2}, 0 < r \le a, \\ 1, r > a. \end{cases}$

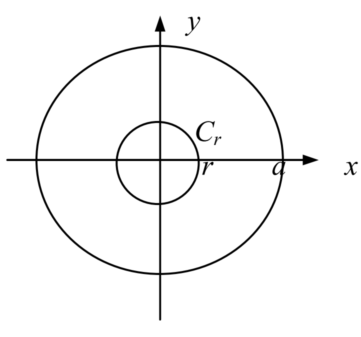


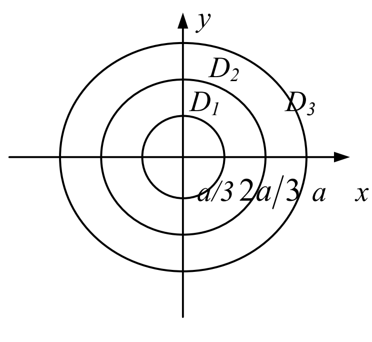
$$F_\xi(z) = P(\xi < z) = \begin{cases} 0, z \le 1, \\ \dfrac{\pi a^2 - \frac{4}{9}\pi a^2}{\pi a^2} = \dfrac{5}{9}; 1 < z \le 5, \\ \dfrac{\pi a^2 - \frac{1}{9}\pi a^2}{\pi a^2} = \dfrac{8}{9}; 5 < z \le 10, \\ 1, z > 10. \end{cases}$$



Зручно вводити простір елементарних подій стандартним способом, при цьому ω={ξ=x}. Результат іспиту – влучення в одну з цих точок.







lx={ξ=x}- влучення

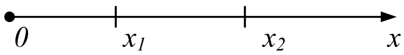
Cr={ρ=r} – коло радіуса r

Dk={ζ=xk}– в область Dk

Властивості функції розподілу

1. $F_\xi(x)$ – неспадна функція.

Хай $x_1<x_2$.



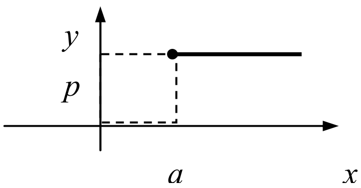
$F_\xi(x_2)=P(\xi<x_2)=P(\xi<x_1)+P(x_1<\xi<x_2)>P(\xi<x_1)=F_\xi(x_1)$; $F_\xi(x_1)<F_\xi(x_2)$;

2. $F_\xi(+\infty)=1$; $F_\xi(+\infty)=P(\xi<\infty)=1$;

$P(-\infty<\xi<\infty)=1$; $F_\xi(-\infty)=0$;

3. $P(\alpha\le\xi<\beta)=P(\xi<\beta)-P(\xi<\alpha)=F_\xi(\beta)-F_\xi(\alpha)$.

Якщо функція розподілу в деякій точці $\xi=a$ зазнає стрибка p, то $P(\xi=a)=p$.



Розглянемо $[a, b)$, $b\to a+0$.

$P(\xi=a)=\lim\limits_{b\to a+0} P(a\le\xi<b)=\lim\limits_{b\to a+0}(F_\xi(b)-F_\xi(a))=p$.

Найбільш важливими типами випадкових величин є дискретні і неперервні випадкові величини.

2.2 Дискретна випадкова величина

Геометричний, біноміальний і пуасонівський закони розподілу

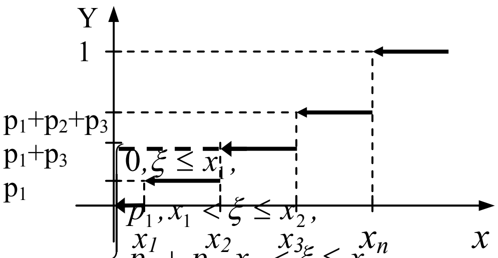
Дискретна випадкова величина

Випадкова величина називається дискретною, якщо вона набуває скінченну або счисленну множину значень (її можливі значення можна перенумерувати).

Нехай $x_1, x_2, \ldots, x_n$ – можливі значення в порядку зростання.

Випадкові події $[\xi=x_1]$, $[\xi=x_2]$, …$[\xi=x_n]$ утворюють повну систему.

$P(\xi=x_k)=p_k$; $\sum_{k=1}^{n} p_k = 1$;



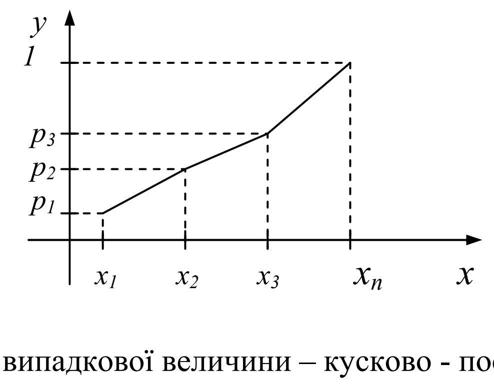
$$F_\xi(x)=P(\xi<x)=\begin{cases} 0, \xi \le x_1, \\ p_1, x_1 < \xi \le x_2, \\ p_1 + p_2, x_2 < \xi \le x_3, \\ .... \\ p_1 + p_2 + ... + p_n = 1, \xi \ge 1. \end{cases}$$

$$F_\xi(x)=\sum_{k=1}^{n} p_k \cdot 1(x - x_k), \qquad де\ 1(x - x_k) = \begin{cases} 1, x \ge x_k, \\ 0, x < x_k. \end{cases}$$

Закон розподілу дискретної випадкової величини можна задати таблицею

ξ	x1	x2	…	xn
P	p1	p2	…	pn

Геометричне завдання дискретної випадкової величини – точки на площині (xi, pi). Ламана, що з'єднує ці точки, називається багатокутником розподілу.



Функція розподілу дискретної випадкової величини – кусково - постійна, у точці xi вона одержує збільшення, що дорівнює ймовірності pi.

$$P(a < \xi < b) = \sum_k p_k, \ x_k \in (a,b).$$

Геометричний закон розподілу

Здійснюється серія випробувань до першої появи події А. Імовірність появи події А в кожному випробуванні дорівнює p і не залежить від результатів інших випробувань.

Нехай ξ - число випробувань.

Подія ξ=k означає, що в (k-1)-ому випробуванні подія А не з'явилася, а в k-ому – з'явилась:

$$p(\xi = k) = p_k = (1-p)^{k-1}p, \qquad k = 0,1,...,n$$.

Позначимо q = 1 - p.

Закон розподілу (ряд розподілу).

ξ	1	2	3	…	k	…	n
P	P	qp	q2p	…	qk-1p	…	qn -1p

$$F\xi(x)= \begin{cases} 0, x \le 1, \\ p \ ,1 < x \le 2, \\ p + pq, 2 < x \le 3, \\ .... \\ p + pq + ... + q^{n-2}p, n-1 < x \le n. \end{cases}$$

$$\sum_{k=1}^{n} p_k = 1.$$

Біноміальний розподіл

Розглядається серія з n випробувань, у кожному з яких подія А чи з’являється, чи не з’являється. Імовірність появи події А постійна і не залежить від результату випробування. Це схема Бернуллі.

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

ξn– випадкова величина, що дорівнює числу появ події А в n випробуваннях.

$$P(\xi_n = k) = C_n^k p^k (1-p)^{n-k}, \qquad k = 0,1,...n$$.

Розподіл називається біноміальним, тому що його коефіцієнти є членами розкладання бінома (p+q)n.

Біноміальний закон розподілу

ξn	0	1	…	k	…	n
pn	qn	npqn-1	…	$C_n^k p^k q^{n-k}$	…	pn

Розподіл Пуассона

Число випробувань необмежено зростає, але так, що np = λ (стала), тобто при великій серії випробувань відбуваються рідкі події.

$$p_k = C_n^k\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1)...(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^n \frac{1}{\left(1-\frac{\lambda}{n}\right)^k} \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!}e^{-\lambda};$$

$$p_k = \frac{\lambda^k}{k!}e^{-\lambda}, \qquad k = 0,1,2...$$

ξ	0	1	…	k	…
P	e-λ	λe-λ	…	$\frac{\lambda^k}{k!}e^{-\lambda}$	…

Число λ називається параметром розподілу.

Розглянемо типову задачу, що приводить до розподілу Пуассона.

Нехай подія А означає відмову складного пристрою протягом малого проміжку часу.

Причиною відмови є вихід з ладу якої-небудь деталі.

Режим роботи пристрою не міняється з часом, відмови окремих деталей не залежать одна від іншої, причому за одиницю часу в «середньому» відбувається λ відмов.

За таких припущень з великим ступенем точності виконуються наступні умови:

Імовірність появи відмови на (0, T) така сама, як і на (t, t+T).

Появи відмов на проміжках часу, які не перетинаються, незалежні.

Ймовірність   появи   відмови за нескінченно малий проміжок часу

p(A) = λ Δt + o(Δt), Δt→0.

Ймовірність появи більш ніж однієї відмови o(Δt), Δt→0.

Розіб'ємо (t, t +T) на n рівних частин $\Delta t = \frac{T}{n.}$.

Розглянемо реєстрацію відмови як окреме випробування:

$$p = \lambda\frac{T}{n} + o(\Delta t); n \to \infty, p \to 0; o(\Delta t) \to 0, \Delta t \to 0,$$

$$np \to \lambda T.$$

Приходимо до розподілу Пуассона для числа відмов за час T:

$$p_T(k)=\frac{(\lambda T)^k}{k!}e^{-\lambda T},k=0,1,...$$

Задачі

2.2.1. Імовірність відмови у роботі пожежного сповіщувача дорівнює p = 0,001. Скільки разів потрібно протестувати один і той же пожежний сповіщувач, щоб з ймовірністю не менше 0,99 він відмовився працювати? (Вважати, що повторні тестування не змінюють p)

$$p+qp+...+q^{n-1}p\geq 0{,}99,$$

$$p\frac{1-q^n}{1-q}=1-q^n\geq 0{,}99,$$

$$1-0{,}999^n\geq 0{,}99,$$

$$0{,}999^n\leq 0{,}01,$$

$$n\lg 0{,}999\leq \lg 0{,}01,$$

$$n\geq\frac{-2}{\lg 0{,}999}\approx 4840.$$

2.2.2. Середнє число пожежних викликів, що надходять на ЦУСЗ в одну хвилину, дорівнює 3. Знайти ймовірність того , що за дві хвилини надійде 4 виклики.

Скористаємося розподілом Пуассона.

$$p_k(t)=\frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

.

P4(2)= $\frac{(3\cdot 2)^4}{4!}e^{-3\cdot 2}=0{,}135.$

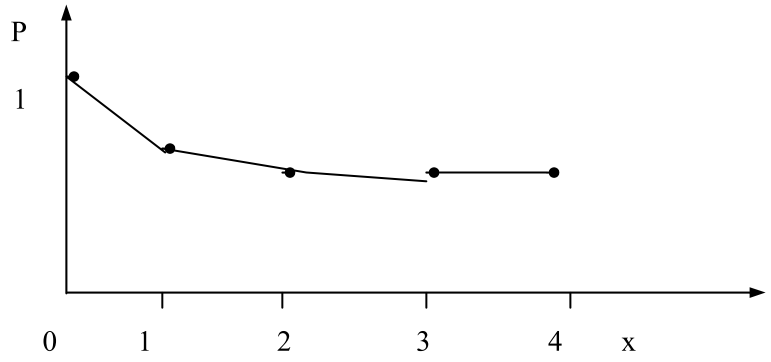
2.2.3. У партії 10% нестандартних деталей. Навмання відібрані 4 деталі. Написати біноміальний закон розподілу дискретної випадкової величини ξ нестандартних деталей серед чотирьох відібраних і побудувати многокутник отриманого розподілу.

ξ	0	1	2	3	4
p	$C_4^0\cdot 0{,}9^4$	$C_4^1 0{,}1\cdot 0{,}9^3$	$C_4^2\cdot 0{,}1^2 0{,}9^2$	$C_4^3 0{,}1^3\cdot 0{,}9$	0,14

Ряд розподілу випадкової величини ξ

ξ	0	1	2	3	4
p	0, 6561	0,2916	0, 04686	0, 0036	0,0001

–



2.2.4. Підручник виданий тиражем 100000 екземплярів. Ймовірність того, що підручник зброшуровано неправильно, дорівнює 0,0001. Знайти ймовірність того, що тираж містить рівно 5 бракованих книг.

У задачі представлена схема рідких подій

P(ξ=5)= $\frac{\lambda^5}{5!}e^{-\lambda} = |\lambda = np = 10| = \frac{10^5}{5!}e^{-10} \approx 0{,}0375$.

2.2.5. По мішені здійснюється залп із 10 гвинтівок. Ймовірність влучення з кожної гвинтівки дорівнює 0,4. Знайти ймовірність хоча б трьох улучень.

Постріл – окреме випробування.

ξ – число влучень у мішень у результаті залпу.

$$p(\xi \ge 3) = 1 - p(\xi \le 2) = 1 - \sum_{k=0}^{2} C_{10}^{k}(0{,}4)^k(0{,}6)^{10-k} =$$

$$= 1 - 0{,}6^{10} - 10 \cdot 0{,}4 \cdot 0{,}6^9 - \frac{10 \cdot 9}{2} \cdot 0{,}4^2 0{,}6^8 \approx 0{,}83.$$

2.2.6. Знайти найімовірніше число правильно набраних слів серед 19, якщо ймовірність того, що слово набрано невірно, дорівнює 0,1.

Число k0 називають найімовірнішим, якщо ймовірність того, що подія відбудеться у цих випробуваннях k0 разів, не менше ймовірності інших можливих результатів випробувань.

$$p_n(k_0) = \max_k p_n(k); \qquad p_n(k) = C_n^k p^k q^{n-k};$$

а) $C_n^{k_0} p^{k_0} q^{n-k_0} \ge C_n^{k_0+1} p^{k_0+1} q^{n-(k_0+1)}$,

б) $C_n^{k_0} p^{k_0} q^{n-k_0} \ge C_n^{k_0-1} p^{k_0-1} q^{n-(k_0-1)}$,

а) $\dfrac{n(n-1)\dots(n-k_0+1)(k_0+1)!\, p^{k_0} q^{n-k_0}}{k_0!\, n(n-1)\dots(n-k_0)p^{k_0-1}q^{n-(k_0+1)}} \ge 1;$

$$\frac{(k_0+1)q}{(n-k_0)p} \geq 1;$$ k0q+q≥np-k0p;

k0≥np-q;

б) $$\frac{n(n-1)...(n-k_0+1)(k_0-1)!\,p^{k_0}q^{n-k_0}}{k_0!\,n(n-1)...(n-k_0+2)p^{k_0-1}q^{n-(k_0-1)}} \geq 1;$$

$$\frac{(n-k_0+1)p}{k_0 q} \geq 1;$$ (n+1)p-k0p≥k0p;

k0≤np+p;

np-q≤k0≤np+p.

19·0,9-0,1≤k0≤19·0,9+0,9;

17≤k0≤18.

Існують два найімовірніших числа: 17 та 18.

2.3 Неперервна випадкова величина. Щільність ймовірності. Рівномірний, показниковий і нормальний розподіли

Неперервна випадкова величина цілком заповнює відрізок.

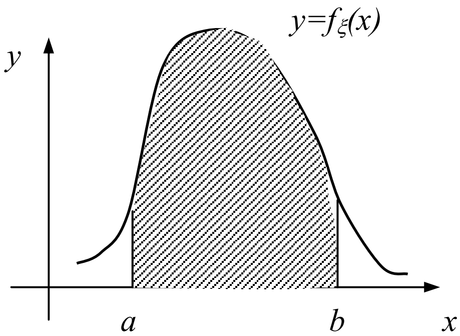
Визначення. Щільністю ймовірності випадкової величини ξ у точці x називається

$$f_\xi(x) = \lim_{\Delta x \to 0} \frac{P(x \leq \xi < x + \Delta x)}{\Delta x}.$$

Властивості.

1. fξ(x) = F/ξ(x).

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



3. fξ(x)≥0 (Fξ(x) неспадна).

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

P(a<ξ<b)= $\int_a^b f_\xi(x)dx.$

Для неперервної випадкової величини P(ξ=x)=0, але

P(x<ξ<x+Δx)≈fξ(x)Δx.

Для неперервної випадкової величини

P(ξ>x)=P(ξ≥x); P(ξ<x)=P(ξ≤x).

Випадкова величина ξ називається неперервною, якщо її розподіл неперервний на всій осі Ох, а щільність ймовірності існує скрізь, за винятком окремих точок.

Функції fξ(x) і Fξ(x) називають диференціальним та інтегральним законами розподілу випадкової величини ξ, відповідно.

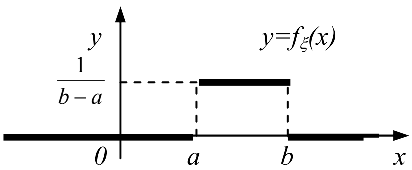
Рівномірний розподіл

Випадкова величина називається рівномірно розподіленою на інтервалі (a, b), якщо

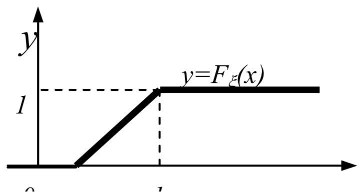
$$f_\xi(x)=\begin{cases} k, & a<x<b, \\ 0, & x<a,\ x>b. \end{cases}$$

$$\int_{-\infty}^{\infty} f_\xi(x)dx=\int_a^b kdx=k(b-a)=1;\quad k=\frac{1}{b-a}.$$

$$f_\xi(x)=\begin{cases} \dfrac{1}{b-a}, & a<x<b, \\ 0, & x<a;\ x<b. \end{cases}$$



$$F_\xi(x)=\begin{cases} \int_{-\infty}^{x} 0dx=0,\ x<a, \\ \int_{-\infty}^{a} 0dx+\int_a^x \dfrac{dx}{b-a}=\dfrac{x-a}{b-a},\ a<x<b, \\ \int_{-\infty}^{a} 0dx+\int_a^b \dfrac{dx}{b-a}+\int_b^x 0dx=1,\quad x<b. \end{cases}$$

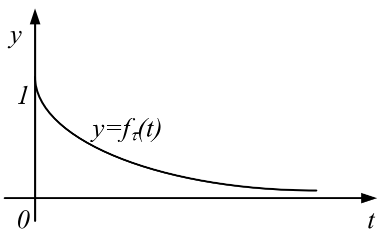


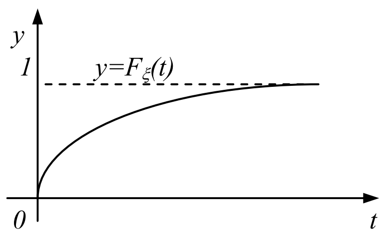
Зміст випадкової величини ξ – помилки округлення обчислень.

Показниковий розподіл. (Експоненціальний закон надійності)

Випадкова величина розподіляється за показниковим законом, якщо її щільність ймовірності має вигляд:

$$f_\tau(t)=\begin{cases} \lambda e^{-\lambda t} & при\ \ t\ge 0, \\ 0, & t<0; \end{cases}\qquad F_\tau(t)=p(\tau<t)=\begin{cases} 1-e^{-\lambda t} & при\ \lambda\ge 0, \\ 0, & \lambda<0. \end{cases}$$

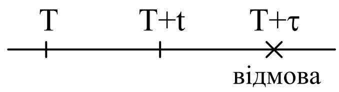




Показниковий розподіл виникає в задачі про відмови (розділ 2.2).

Нехай τ – проміжок часу від довільного фіксованого моменту Т до найближчої відмови.

Кількість відмов за час t > 0 підкоряється закону Пуассона з параметром λt.



Оскільки подія {τ > t} означає відсутність відмов на інтервалі (T, T+t) (за час t), то

$$p(\tau > t) = p_t\,(k=0) = \left.\frac{(\lambda t)^k}{k!}e^{-\lambda t}\right|_{k=0} = e^{-\lambda t}.$$

$$F_\tau(t) = 1 - P(\tau > t) = \begin{cases} 1 - e^{-\lambda r}, t \ge 0, \\ 0, \qquad t > 0. \end{cases}$$

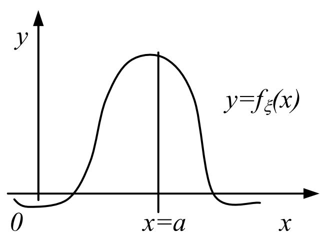
Зміст випадкової величини τ – час безвідмовної роботи елемента (системи).

Нормальний розподіл. (Закон Гаусса)

Виникає в результаті накладення великого числа незалежних чи слабко залежних впливів.

Диференційна функція розподілу (щільність імовірностей) нормально розподіленої випадкової величини ξ дорівнює

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



де a – середнє значення;

x = a – вісь симетрії функції $f_\xi(x)$;

σ – середнє квадратичне відхилення, що характеризує ширину кривої відносно x = a.

Інтегральна функція розподілу

$$F_\xi(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$
.

Fξ(x) у загальному виді не обчислюється, але виражається через функцію Лапласа Φ(x):

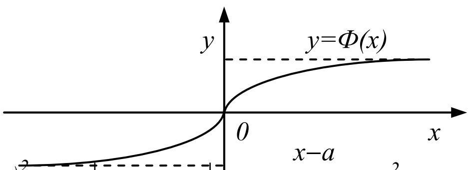
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt.$$

Властивості Φ(x):

1. Φ(-x)=-Φ(x);   Φ(0)=0,

2. Φ(∞)= $\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{t^2}{2}}dt=\left|\begin{matrix}\frac{t}{\sqrt{2}}=x\\ dt=\sqrt{2}dx\end{matrix}\right|=\frac{\sqrt{2}}{\sqrt{2\pi}}\int_0^\infty e^{-x^2}dx=\frac{1}{\sqrt{\pi}}\cdot\frac{\sqrt{\pi}}{2}=\frac{1}{2};$

Φ(-∞)=-1/2.



$$F_\xi(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}}dt=\left|\begin{matrix}\frac{t-a}{\sigma}=u\\ dt=\sigma du\end{matrix}\right|=\frac{1}{2\pi}\int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{u^2}{2}}du=$$

$$=\frac{1}{2\pi}\int_{-\infty}^{0} e^{-\frac{u^2}{2}}du+\frac{1}{\sqrt{2\pi}}\int_{0}^{\frac{x-a}{\sigma}} e^{-\frac{u^2}{2}}du=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right),$$

$$F_\xi(x)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right),$$

$$P(x_1<\xi<x_2)=\Phi\left(\frac{x_2-a}{\sigma}\right)-\Phi\left(\frac{x_1-a}{\sigma}\right),$$

$$P(|\xi-a|<\varepsilon)=P(a-\varepsilon<\xi<a+\varepsilon)=2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

Випадкову величину ξ інтерпретують як помилки вимірювання величини a.

Знайдемо ймовірність виконання нерівності $P(|\xi\text{-}a| < 3\sigma)$:

$P(|\xi\text{-}a|<3\sigma)=2\Phi(3)\approx 0{,}9972$,

тобто виконання такої нерівності практично вірогідно.

У цьому полягає «правило трьох σ».

Визначення: Модою xMo неперервної випадкової величини називається точка, що задовольняє умову

$$\max_x f_\xi(x) = f(x_{Mo}).$$

Медіаною xMe неперервної випадкової величини називається корінь рівняння F(x)=1/2.

Для нормального закону розподілу точка a одночасно є і модою, і медіаною: a=xMo=xMe.

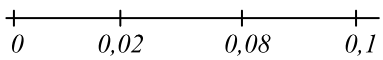
Задачі

2.3.1. Ціна поділки шкали амперметра дорівнює 0,1 А. Показання округлюють до найближчої поділки. Знайти ймовірність того, що при відліку буде зроблена помилка, що перевищує 0,02.

$P(|\xi|>0{,}02)$

$f\xi(x)=1/0{,}1 = 10$ – щільність ймовірності похибок округлення.

$P(0{,}02<\xi<0{,}08)=\int_{0,02}^{0,08} 10dx = 10x\Big|_{0,02}^{0,08} = 0{,}6$



2.3.2. Відбуваються порівняльні випробування двох незалежно працюючих водяних насосів. Тривалість часу безвідмовної роботи першого має розподіл:

p1(τ<t) = F1(t)=1-e-0,02t,

а другого

p2(τ<t) = F2(t)=1-e-0,05t.

Знайти ймовірність того, що за тривалість часу t = 6

а) відмовлять обидва насоси;

б) обидва насоси не відмовлять;

в) відмовить хоча б один насос.

Тривалість часу безвідмовної роботи буде t, якщо t < τ, а p(τ > t) = p = e-λt.

Ймовірності безвідмовної роботи за час t для першого і другого насосів:

p1=p(τ>6)=e-0,12,

p2=p(τ>6)=e-0,3.

Ймовірність того, що відмовлять перший і другий насоси:

q1=1-p1=1-e- 0,12 q2=1-p2=1-e- 0,3.

Ймовірність того, що

а) обидва насоси відмовлять за час t:

p=q1q2=(1-e-0,12)(1-e-0,3);

б) тільки один насос відмовить:

p=p1q1+p2q2=e-0,12 (1-e-0,12)+e-0,3(1-e-0,3);

в) хоча б один насос відмовить:

p=1-p1p2-=1-e-0,12e-0,3.

2.3.3. Автомат виготовляє кульки. Кулька вважається придатною, якщо відхилення X діаметра кульки від проектного розміру a за абсолютною величиною менше 0,7 мм. Вважаючи, що випадкова величина X розподілена нормально із середнім квадратичним відхиленням σ = 0,4 мм, знайти скільки буде придатних кульок серед ста виготовлених.

Скористаємося формулою

$$p(|x-a|<\varepsilon)=2\Phi\left(\frac{\varepsilon}{\sigma}\right)=2\Phi\left(\frac{0{,}7}{0{,}4}\right)=2\Phi(1{,}75)=2\cdot 0{,}4599=0{,}92,$$

звідси випливає, що приблизно 92 кульки з 100 виявляться придатними.

2.3.4. Випадкова величина ξ в інтервалі (3, 5) задана диференціальною функцією

$$f_\xi(x)=-\frac{3}{4}x^2+6x-\frac{45}{4},$$

поза цим інтервалом f(x) = 0. Знайти моду і медіану ξ.

Знайдемо моду з умови $\max\limits_{x\in(3,5)} f_\xi(x)$.

$f'_\xi(x)=-\frac{3}{2}x+6=0; x=4$. Оскільки $f''_\xi(x)=-\frac{3}{2}<0$, то $\max\limits_{x\in(3,5)} f_\xi(x)=f_\xi(4)$

таким чином, xM0 = 4.

$$f_\xi(3)=-\frac{3}{4}\cdot 3^2+6\cdot 3-\frac{45}{4}=\frac{-27+72-45}{4}=0,$$

$$f_\xi(5)=-\frac{3}{4}\cdot 5^2+6\cdot 5-\frac{45}{4}=\frac{-75+120-45}{4}=0,$$

отже, функція $f_\xi(x)$ - симетрична відносно прямої x = 4; площа області

$$y=-\frac{3}{2}x+6x-\frac{45}{4},\qquad y=0,\qquad 3\le x\le 5$$

у точці x = 4 поділяється навпіл, тому xMe =4.

2.4 Двовимірний випадковий вектор

Нехай Ω – простір елементарних подій.

Випадковим вектором (ξ,η) називається пара числових функцій елементарної події ξ=f1(ω) η=f2(ω), для яких при будь-яких значеннях x,y існує P(ξ < x, η < y).

З випадковим вектором (ξ, η) можна зв'язати простір елементарних подій виду [ξ = x, η = y], де (x, y) – будь-яка можлива подія вектора (ξ, η). Геометрично події зображуються точками площини (x, y), а результат випробування – влученням в одну з цих точок. Ω збігається з площиною x, y чи її частиною.

$F_{\xi\eta}(x, y) = p(\xi < x, \eta < y)$ – функція розподілу випадкового вектора (ξ, η).

Визначення: Вектор називається дискретним, якщо його можливі значення можна перенумерувати, тобто (ξ,η)= $\{(x_i, y_j), i = \overline{1,n}; j = \overline{1,m}\}$

Закон розподілу:

P(ξ =xi, η = yi) = Pij.

Ряд розподілу задається двовимірною таблицею

η \ ξ	x1	x2	…	xn
y1	P11	P12	…	P1n
y2	P21	P22	…	P2n
…	…	…	…	…
ym	Pm1	Pm2	…	Pmn

P(ξ =xi) = Pi = $\sum_{j=1}^{m} P_{ij}$;

P(η = yj) = Pj = $\sum_{i=1}^{n} P_{ij}$;

Умовні закони розподілу

$$P(\xi = x_i / \eta = y_j) = \frac{P(\xi = x_i; \eta = y_j)}{P(\eta = y_j)} = \frac{P_{ij}}{P_j};$$

$$P(\eta = y_j / \xi = x_i) = \frac{P(\eta = y_j; \xi = x_i)}{P(\xi = x_i)} = \frac{P_{ij}}{P_i};$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} = 1$$ (за умови нормування).

$$\sum_{i=1}^{n} P_i = \sum_{j=1}^{m} P_j = 1.$$

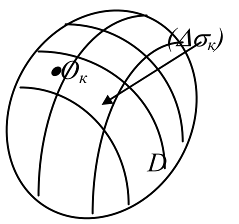
Щільністю ймовірності fξ,η (x,y) випадкового вектора (ξ, η) у точці θ(x,y) називається границя відношення ймовірності влучення в область G_k, що містить точку θ, до площі $\Delta\sigma_k$ області G_k, якщо область G_k стягається в точку θ:

$$f_{\xi,\eta}(x,y)=\lim_{\Delta G_k\to\theta}\frac{P[(\xi,\eta)\in\Delta G_k}{\Delta\sigma_k}.$$

Оскільки границя однозначне поняття, то

$$f_{\xi,\eta}(x,y)=\lim_{\substack{\Delta x\to\theta\\ \Delta y\to 0}}\frac{P\{x<\xi<x+\Delta x,y<\eta<y+\Delta y\}}{\Delta x\Delta y}.$$

Випадковий вектор (ξ,η) називається неперервним, якщо для нього ймовірність влучення в будь-яку точку області дорівнює нулю, а щільність ймовірності існує усюди, за винятком, може бути , окремих точок.



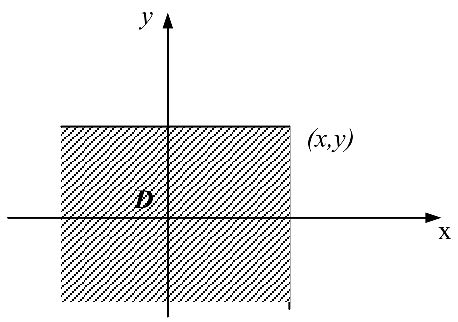
$$P((\xi,\eta)\in D)\approx\sum_{k=1}^{n}f_{\xi,\eta}(\theta_k)\Delta\sigma_k+o(\Delta\sigma_k)\xrightarrow[d(\Delta\sigma_k)\to 0]{n\to\infty}\iint_D f_{\xi,\eta}(x,y)d\sigma.$$

$$P((\xi,\eta)\in D)=\iint_D f_{\xi,\eta}(x,y)dxdy.$$

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

$$f_{\xi,\eta}(x,y)=\frac{\partial^2F(x,y)}{\partial x\partial y},$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{\xi,\eta}(x,y)dxdy=1.$$



Рівномірний розподіл на площині

Випадковий вектор (ξ,η) називається рівномірно розподіленим в області D, якщо

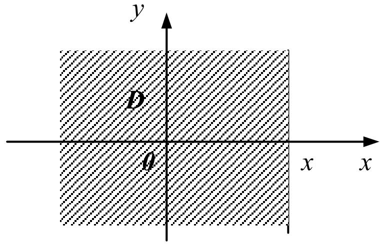
$$f_{\xi,\eta}(x,y)=\begin{cases}k,(x,y)\in D;\\0,(x,y)\in\bar{D}.\end{cases}1=\iint_D kd\sigma=k\sigma_D;k=\frac{1}{\sigma_D}.$$

Закони розподілу проекцій випадкового вектора

$$F_{\xi,\eta}(x,y) = P(\xi < x, -\infty < \eta < \infty) = \int_{-\infty}^{x} dx \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy,$$

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

$$f_{\eta}(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx.$$



Проекції неперервного випадкового вектора є неперервними випадковими величинами. Зворотна задача відновлення fξ,η(x,y) по fξ(x) і fη(y) у загальному випадку нерозв'язна, тому що невідомий взаємозв'язок між ξ і η.

Визначення. Умовною щільністю ймовірності fξ(x/η = y) називається

$$f_{\xi}(x|_{\eta=y}) = f(x/y) = \lim_{\Delta x \to 0} \frac{P(x < \xi < x + \Delta x / \eta = y)}{\Delta x}$$

$$f_{\xi}(x|_{\eta=y}) = \lim_{\Delta x \to 0} \frac{P(x < \xi < x + \Delta x|_{y<\eta<y+\Delta y})}{\Delta x} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P(x < \xi < x + \Delta x, y < \eta < y + \Delta y)}{\Delta x \Delta y P(y < \eta < y + \Delta y)/\Delta y} =$$

$$= \frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}.$$

$$f_{\xi}(x|_{\eta=y}) = \frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}; \qquad f_{\eta}(y|_{\xi=x}) = \frac{f_{\xi,\eta}(x,y)}{f_{\xi}(x)}.$$

$$\int_{-\infty}^{\infty} f_{\xi}(x|_{\eta=y})dx = \frac{1}{f_{\eta}(y)} \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx = 1,$$

$$\int_{-\infty}^{\infty} f_{\eta}(y|_{\xi=x})dy = \frac{1}{f_{\xi}(x)} \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy = 1$$

$$f_{\xi,\eta}(x,y) = f_{\xi}(x|_{\eta=y})f_{\eta}(y) = f_{\eta}(y|_{\xi=x})f_{\xi}(x) \qquad (2.4.1)$$

З (2.4.1) можна одержати формули, аналогічні до формул Байєса

$$f_{\xi}(x/y) = \frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}; f_{\eta}(y/x) = \frac{f_{\xi,\eta}(x,y)}{f_{\xi}(x)}.$$

Визначення. Випадкові величини ξ, η називаються незалежними, якщо при будь-яких можливих значеннях x, y їхні умовні щільності ймовірності збігаються з безумовними

$$f_\xi(x/y)=f_\xi(x);\, f_\eta(y/x)=f_\eta(y)$$,

чи

$$f_{\xi,\eta}(x,y)=f_\xi(x)f_\eta(y).$$

Ознаки незалежності випадкових величин

Якщо двовимірна щільність розподілу fξ,η(x,y) представляється у виді

$$f_{\xi,\eta}(x,y)=\alpha(x)\beta(y)$$,

то ξ,η – незалежні.

$$f_{\xi,\eta}(x,y)=\alpha(x)\beta(y)\;;\quad 1=\int_{-\infty}^{\infty}\alpha(x)dx\int_{-\infty}^{\infty}\beta(y)dy.$$

$$f_\xi(x)=\alpha(x)\int_{-\infty}^{\infty}\beta(y)dy=\frac{\alpha(x)}{\int_{-\infty}^{\infty}\alpha(x)dx};\qquad f_\eta(y)=\frac{\beta(y)}{\int_{-\infty}^{\infty}\beta(y)dy}.$$

Звідси випливає, що

$$f_{\xi,\eta}(x,y)=f_\xi(x)f_\eta(y).$$

Задачі

2.4.1. Задано дискретну двовимірну випадкову величину (ξ,η)

η \ ζ	x1=2	x2=5	x3=8
y1=0,4	0,15	0,30	0,35
y2=0,8	0,05	0,12	0,03

Знайти :

а) безумовні закони розподілу складових;

б) безумовний закон розподілу складової ξ за умови, що η прийняла значення y1 = 0,4;

в) умовний закон розподілу за умови, що ξ = x2 = 5.

а) p(ξ =xi; η = (y1,y2)) = $\sum_{k=1}^{2}p_{ik};\quad i=\overline{1,3}.$

ξ	2	5	8
p	0,20	0,42	0,38

p(ξ=xi,x2,x3);η=yk) = $\sum_{k=1}^{2} p_{ik};$

η	0,4	0,8
p	0,80	0,20

б) p(ξ=xi/η=y1)=

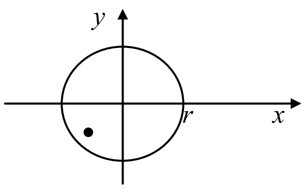
ξ	2	5	8
P(ξi/η=y1)	3/16	3/8	7/16

в) p(η=yk/ξ=x2)= $\frac{p(\xi = x_2, \eta = y_k)}{p(\xi = x_2)} = \frac{p(\xi = x_2, \eta = y_k)}{0{,}42}; \quad k = 1{,}2.$

η	0,4	0,8
p(η/x2)	$\frac{0{,}30}{0{,}42} = \frac{5}{7}$	$\frac{0{,}12}{0{,}42} = \frac{3}{7}$

2.4.2. Неперервна випадкова величина (ξ,η) розподілена рівномірно в колі радіуса r з центром на початку координат. Довести, що ξ і η залежні.

За умови, $f_{\xi\eta}(x,y) = \frac{1}{\pi r^2}$ в колі радіуса r і $f_{\xi\eta}(x,y) = 0$ поза колом.



$$f_\xi(x) = \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\pi r^2} dy = \frac{2}{\pi r^2}\sqrt{r^2 - x^2}, \qquad |x| \le r;$$

$$f_\eta(y) = \frac{2}{\pi r^2}\sqrt{r^2 - y^2}; \qquad |y| \le r^2;$$

$$f(x/y) = \frac{f_{\xi,\eta}(x,y)}{f_\eta(y)} = \frac{1}{2\sqrt{r^2 - y^2}}; \qquad f(y/x) = \frac{f_{\xi,\eta}(x,y)}{f_\xi(y)} = \frac{1}{2\sqrt{r^2 - x^2}},$$

отже ξ і η - залежні.

2.4.3. Випадковий вектор (ξ, η) має щільність ймовірності

$$f_{\xi,\eta}(x,y)=\frac{1}{\pi^2(1+x^2+y^2+x^2y^2)}.$$

Чи є випадкові величини ξ і η залежними?

$$f_{\xi,\eta}(x,y)=\frac{1}{\pi^2(1+x^2)(1+y^2)}=\alpha(x)\cdot\beta(y),$$

отже ξ і η незалежні.

2.5 n-мірний випадковий вектор. Функції випадкових величин

n-мірний випадковий вектор

Хай Ω – простір елементарних подій.

Визначення. Випадковим вектором $\vec{\xi}=(\xi_1,\xi_2,...\xi_n)$ називається сукупність числових функцій елементарної події ω

ξ1=φ1(ω); ξ2=φ2(ω);…;ξn=φn(ω),

для якої існує P(ξ1 < x1, ξ2 < x2,…,ξn < xn)...

F(x1,x2,…xn)=P(ξ1 < x1, ξ2 <x2,…,ξn<xn)–функція розподілу випадкової величини $\vec{\xi}$.

$\vec{x}=(x_1,x_2,...x_n)$ – можливі значення випадкового вектора, тобто точки n-мірного простору R^n. Результат випробування – влучення в точку R^n.

ω = {ξ1= x1,…, ξn = xn} – елементарна подія.

$$f_{\vec{\xi}}(x_1,x_2,...,x_n)=\lim_{\substack{\beta_i-\alpha_i\to 0\\ i=1,n\\ \alpha_i<x_i<\beta_i}}\frac{P(\alpha_1<\xi_1<\beta_1,...\alpha_n<\xi_n<\beta_n)}{(\beta_1-\alpha_1)...(\beta_n-\alpha_n)},$$

$$P(\vec{\xi}\in D)=\int\limits_D\!...\!\int f_{\vec{\xi}}(x_1,x_2,...x_n)dx_1dx_2...dx_n.$$

$$f_{\xi_1;\xi_2;...\xi_{n-1}}(x_1,x_2,...,x_{n-1})=\int_{-\infty}^{\infty}f_{\vec{\xi}}(x_1,x_2,...x_n)dx_n$$ – щільність ймовірності вектора $\xi_1,\xi_2,...,\xi_{n-1}$.

$$f_{\xi_1}(x_1/\xi_2 = x_2,...,\xi_n = x_n) = \frac{f_{\xi_1,...\xi_n}(x_1, x_2,..., x_n)}{f_{\xi_2,...\xi_n}(x_2, x_3,..., x_n)},$$

$$f_{\xi_1,\xi_2}(x_1, x_2/\xi_3 = x_3,...,\xi_n = x_n) = \frac{f_{\vec{\xi}}(x_1, x_2,..., x_n)}{f_{\xi_3,\xi_4,...\xi_n}(x_3, x_4,..., x_n)}.$$

Для взаємної незалежності випадкових величин ξ1, ξ2…ξn необхідно і достатньо, щоб

$$f_{\vec{\xi}}(x_1, x_2,...x_n) = a_1(x_1)...a_k(x_n),$$

$$f_{\xi_k}(x_k) = \frac{\alpha_k(x_k)}{\int_{-\infty}^{\infty} \alpha_k(x_k)dx_k}.$$

Функції випадкових величин

Нехай на Ω задана випадкова величина

$\xi = f(\omega)$.

Під функцією $\eta=\alpha(\xi)$ розуміється нова випадкова величина

$\eta = \alpha[f(\omega)]$.

Таким чином, значення x, y випадкових величин ξ і $\eta = \alpha(\xi)$, що відповідають тому самому ω (тобто одержані в результаті того самого випробування), зв'язані функціональною залежністю

$Y = \alpha(x)$.

Приклад 1. Знаючи закон розподілу випадкової величини ξ, знайти закон розподілу випадкової величини $\eta = \alpha(\xi)$.

Нехай $\xi = \{x_i\}_1^n ; \eta = \{\alpha(x_i)\}_1^n$

ξ	x1	x2	…	xn
	p1	p2	…	pn

Якщо всі α(xk) - різні, то події [ξ = xk] і [η = α(xk)] - тотожні.

η	α(x1)	α(x2)	…	α(xn)
p	p1	p2	…	pn

Якщо серед чисел α(xk) є однакові, тоді кожній групі однакових значень слід відвести в таблиці один стовпець, а їхні ймовірності додати.

Приклад 2. P(ξk = k) = $C_n^k p^k q^{n-k}$; ξk = 0,1,…k,…n.

$$\eta_n = \frac{\xi_n}{n}.$$

Знайти P($\eta_k = \frac{k}{n}$).

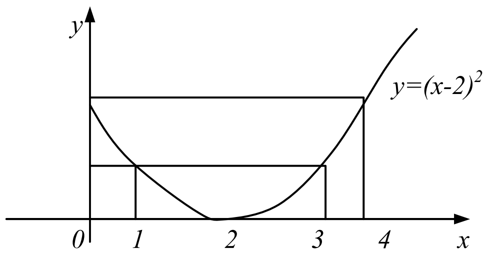
P($\eta_k = \frac{k}{n}$)=P($\xi_k = k$).

η	0	$\frac{1}{n}$	…	$\frac{k}{n}$	…	1
p	qn	npqn-1	…	$C_n^k p^k q^{n-k}$	…	pn

Приклад 3.

$$p(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}; k = 0,1,2,...$$

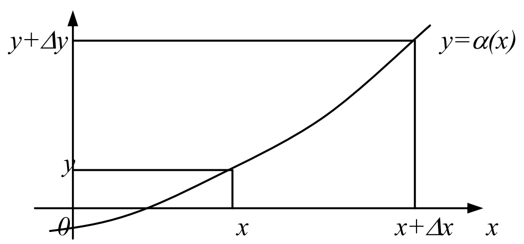
$$\eta = (\xi - 2)^2.$$



Для неперервної випадкової величини задача ставиться так: знаючи щільність ймовірності випадкової величини ξ, знайти щільність ймовірності випадкової величини η = α(ξ).

а) Нехай y = α(x) зростає і є диференційованою.

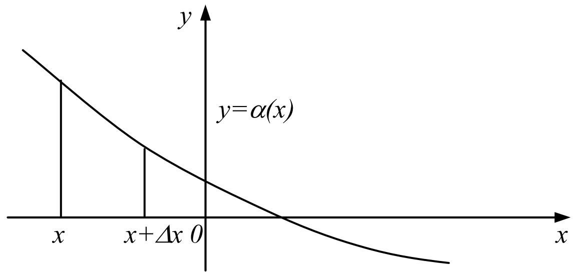
Хай [x<ξ<x+Δx]=[y<η<y+Δy] ,



$$f_\eta(y) = \lim_{\Delta y \to 0} \frac{p(y < \eta < y + \Delta y)}{\Delta y} = \lim_{\substack{\Delta y \to 0 \\ (\Delta x \to 0)}} \frac{p(x < \xi < x + \Delta x)\Delta x}{\Delta y \cdot \Delta x} = f_\xi(x)x'(y) = f_\xi(x)\frac{1}{\alpha'(x)}.$$

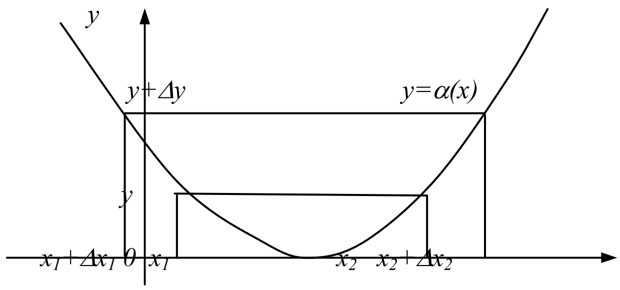
Якщо функція монотонно спадає, то $\Delta x < 0$

$$f_\eta(y) = f_\xi(x)\frac{1}{\left|\alpha^{/}(x)\right|}.$$



Нехай функція $y = \alpha(x)$ має один екстремум .

Хай $x1 = \beta1(y)$; $x2 = \beta2(y)$



$$p(y < \eta < y + \Delta y) = p(x_1 < \xi < x_1 + \Delta x_1) + p(x_2 < \xi < x_2 + \Delta x_2),$$

$$f_\eta(y) = f_\xi(x_1)\frac{1}{\left|\alpha^{/}(x_1)\right|} + f_\xi(x_2)\frac{1}{\left|\alpha^{/}(x_2)\right|},$$

$$f_\eta(y) = \sum_{k=1}^{n} f_\xi(x_k)\left|x_k^{/}(y)\right|.$$

Задачі

2.5.1. Дискретна випадкова величина Х задана законом розподілу

X	3	6	10
P	0,2	0,1	0,7

Знайти закон розподілу випадкової величини $y = 2x + 1$.

Рішення. Обчислюючи значення $yi = 2xi + 1$ для можливих значень випадкової величини Х, дістанемо

Y	7	13	21
P	0,2	0,1	0,7

2.5.2. Дискретна випадкова величина X задана законом розподілу

X	π/4	π/2	3π/4
P	0,2	0,7	0,1

Знайти закон розподілу випадкової величини $Y = \sin X$.

Рішення. Обчислюючи значення $y_i = \sin x_i$ для можливих значень випадкової величини X, дістанемо

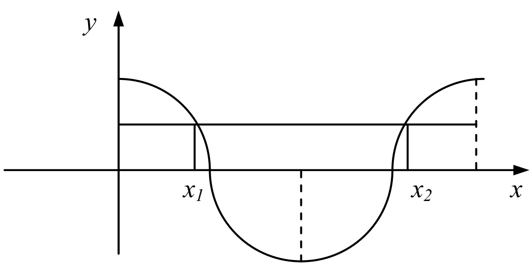
Y	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$
P	0,2	0,7	0,1

Підсумовуючи ймовірності для значень $y_i = \frac{\sqrt{2}}{2}$, дістанемо

Y	$\frac{\sqrt{2}}{2}$	1
P	0,3	0,7

2.5.3. Випадкова величина X розподілена рівномірно в інтервалі (0,2π). Знайти диференціальну функцію g(y) випадкової величини $Y = \cos X$.

Рішення. На інтервалі (0,2π) обернена функція $x = \arccos y$ двозначна (рис. 7.1).



$x_1 = \arccos y$, $x_2 = \pi - \arccos y$,

$$x_1'(y) = -\frac{1}{\sqrt{1-y^2}}, \quad x_2'(y) = \frac{1}{\sqrt{1-y^2}}.$$

Скористаємося формулою

$$f_y(y)=f_x(x_1(y))\,|\,x_1'(y)\,|+f_x(x_2(y))\,|\,x_2'y\,|=\frac{1}{2\pi}\cdot\frac{1}{\sqrt{1-y^2}}+\frac{1}{2\pi}\cdot\frac{1}{\sqrt{1-y^2}}=\frac{1}{\pi\sqrt{1-y^2}}$$

Отже, fy(y)= $\begin{cases}\dfrac{1}{\pi\sqrt{1-y^2}}, -1<y<1,\\ 0, \quad y\le -1,\, y\ge 1.\end{cases}$

2.6 Функціональні перетворення випадкових векторів.

Композиція випадкових векторів

Функціональні перетворення випадкових векторів

Нехай неп $\vec{\xi}=(\xi_1,\xi_2)$ і $\vec{\eta}=(\eta_1,\eta_2)$ – випадкові вектори.

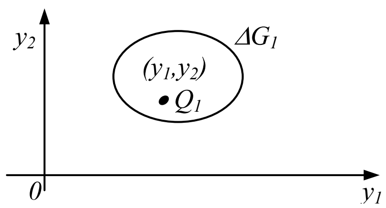
$$\begin{cases}\eta_1=\alpha_1(\xi_1,\xi_2),\\ \eta_2=\alpha_2(\xi_1,\xi_2).\end{cases}$$

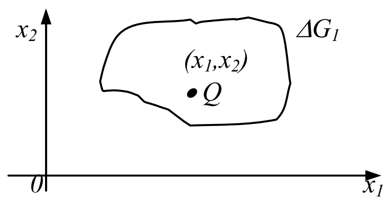
$\vec{\xi}$ – ерервний вектор. $\begin{cases}y_1=\alpha_1(x_1,x_2),\\ y_2=\alpha_2(x_1,x_2).\end{cases}$

Вважаємо, що перетворення взаємно однозначні і функції є диференційованими (x1,x2)↔(y1,y2).

Знайти щільність ймовірності $f_{\vec{\eta}}(y_1,y_2)$ за заданою щільністю ймовірності $f_{\vec{\xi}}(x_1,x_2)$.

$$\begin{cases}x_1=\beta_1(y_1,y_2),\\ x_2=\beta_2(y_1,y_2).\end{cases}$$





$$p\{(\xi_1,\xi_2)\in\Delta G_2\}=p\{(\eta_1,\eta_2)\in\Delta G_1\},$$

$$f_{\vec{\eta}}(y_1,y_2)=\lim_{\Delta G_1\to\theta_1}\frac{p\{(\eta_1,\eta_2)\in\Delta G_1\}}{\Delta\sigma_1}=\lim_{\substack{\Delta G_2\to\theta_1\\(\Delta G_1\to\theta)}}\frac{p\{(\xi_1,\xi_2)\in\Delta G_2\}}{\Delta\sigma_2}\frac{\Delta\sigma_2}{\Delta\sigma_1}=f_{\vec{\xi}}(x_1,x_2)\lim_{\Delta\sigma_1}\frac{\Delta\sigma_2}{\Delta\sigma_1}.$$

$\lim\limits_{\Delta\sigma_1\to 0}\dfrac{\Delta\sigma_2}{\Delta\sigma_1}=\dfrac{\partial(x_1,x_2)}{\partial(y_1,y_2)}$ – якобіан перетворення в точці θ1.

$$f_{\vec{\eta}}(y_1,y_2)=f_{\vec{\xi}}(x_1,x_2)\left|\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}\right| \qquad (2.6.1)$$

Формула (2.6.1) допускає узагальнення на n –вимірний випадок.

Лінійне перетворення випадкового вектора

$$\begin{cases}\eta_1=a_1\xi_1+b_1\xi_2,\\ \eta_2=a_2\xi_1+b_2\xi_2,\end{cases} \qquad \begin{cases}y_1=a_1x_1+b_1x_2,\\ y_2=a_2x_1+b_2x_2,\end{cases} \qquad \Delta=\begin{vmatrix}a_1 b_1\\ a_2 b_2\end{vmatrix},$$

$$x_1=\frac{1}{\Delta}\begin{vmatrix}y_1 b_1\\ y_2 b_2\end{vmatrix}=\frac{1}{\Delta}(y_1b_2-y_2b_1), \qquad x_2=\frac{1}{\Delta}\begin{vmatrix}a_1 y_1\\ a_2 y_2\end{vmatrix}=\frac{1}{\Delta}(y_2a_1-y_1a_2),$$

$$\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}=\begin{vmatrix}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1}\\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2}\end{vmatrix}=\frac{1}{\Delta^2}\begin{vmatrix}b_2 & -a_2\\ -b_1 & a_1\end{vmatrix}=\frac{1}{\Delta^2}(a_1b_2-b_1a_2)=\frac{1}{\Delta},$$

$$f_{\vec{\eta}}(y_1,y_2)=\frac{1}{|\Delta|}f_{\vec{\xi}}=\left(\frac{y_1b_2-y_2b_1}{\Delta},\frac{y_2a_1-y_1a_2}{\Delta}\right).$$

Композиція випадкових величин

Знайти щільність ймовірності суми випадкових величин ξ1, ξ2 за заданою щільністю ймовірності $f_{\vec{\xi}}(\xi_1,\xi_2).$

Розглянемо систему

$$\begin{cases}\eta_1=\xi_1+\xi_2,\\ \eta_2=\xi_2.\end{cases} \quad \frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}=\begin{vmatrix}1 & 0\\ -1 & 1\end{vmatrix}=1$$

$$\begin{cases}y_1=x_1+x_2,\\ y_2=x_2.\end{cases} \quad \begin{cases}x_1=y_1-y_2,\\ x_2=y_2.\end{cases}$$

$$f_{(\eta_1,\eta_2)}(y_1,y_2)=f_{(\xi_1,\xi_2)}(y_1-y_2,y_2),$$

$$f_{\eta_1}(y_1)=\int_{-\infty}^{\infty}f_{(\xi_1,\xi_2)}(y_1-y_2,y_2)dy_2. \qquad (2.6.2)$$

Якщо $\xi 1$ і $\xi 2$ – незалежні, то $$f_{\eta_1}(y)=\int_{-\infty}^{\infty} f_{\xi_1}(y_1-y_2)f_{\xi_2}(y_2)dy_2 \qquad (2.6.3)$$

Задачі

2.6.1. Знайти щільність ймовірності суми двох незалежних випадкових величин, розподілених рівномірно на інтервалі (0, 1).

Рішення. Позначимо випадкові величини X і Y, а Z = X + Y, причому fy(y) = 1, y∈(0,1), fx(x) = 1, якщо x∈(0,1). fy(y) = fx(x) = 0 поза інтервалами (0,1). Оскільки X і Y - незалежні випадкові величини, то

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy .$$

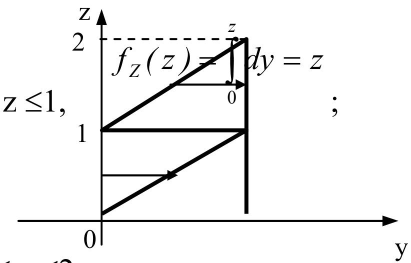
Область інтегрування визначена умовами:

0≤ y ≤1,

0≤ z-y ≤1,

і обмежена прямими: y = 0; y = 1 ; z = y; z = y +1 (рис.2.6.1).

1) 0≤ z ≤1, $f_Z(z)=\int_0^z dy = z$;



2) 1≤ z ≤2, $f_Z(z)=\int_{z-1}^{1} dy = 2-z$.

Рис. 2.6.1

Остаточно маємо

$$f_2(z)=\begin{cases} 0, & z<0, \\ z, & 0\le z<1, \\ 2-z, & 1\le z<2 \\ 0, & z\ge 2. \end{cases}$$

2.6.2. Дискретні незалежні випадкові величини X і Y задані розподілами:

X	1	3
p	0,3	0,7

Y	2	4
p	0,6	0,4

Знайти розподіл випадкової величини Z =X +Y.

Знайдемо можливі значення випадкової величини Z:

z1= x1 +y1 = 1+2=3; p(Z=3)= p(X=1)· p(Y=2)= 0,3·0,6= 0,18,

z2= x1 +y2 = 1+4=5; p(Z=5)= p(X=1)· p(Y=4)= 0,3·0,4= 0,12,

z3 =x2 +y1 = 3+2=5; p(Z=5)= p(X=3)· p(Y=2)= 0,7·0,6= 0,42,

z4 =x2+y2 = 3+4=7; p(Z=7)= p(X=3)· p(Y=4)= 0,7·0,4= 0,28.

p(Z=5)= p(Z=Z1)·+p(Z=Z2)= 0,12+0,42= 0,54.

Розподіл випадкової величини Z:

Z	3	5	7
P	0,18	0,54	0,28

2.7 Числові характеристики випадкових величин. Математичне сподівання

Числові характеристики служать для наближеного опису випадкової величини. Найважливішими є середнє значення і дисперсія.

Середнє значення (математичне сподівання)

а) ξ – дискретна випадкова величина.

Розподіл ймовірності випадкової величини ξ:

ξ	x1	x2	…	xn
p	p1	p2	…	pn

Зробимо серію N незалежних випробувань і позначимо mk число випробувань, у яких випадкова величина ξ набуває значення xk.

$$\xi_N = \frac{1}{N}(x_1 m_1 + \ldots + x_n m_n) = \sum_{k=1}^{n} x_k p_k^* \quad –$$

середнє арифметичне результатів серії випробувань.

Оскільки за великої кількості випробувань pk* ≈ p, то

$$\xi_N \approx \sum_{k=1}^{n} x_k p_k \, .$$

Визначення. Середнім значенням (чи математичним сподіванням) дискретної випадкової величини ξ називається сума добутків її можливих значень на відповідні ймовірності:

$$M\xi = \sum_{k=1}^{n} x_k p_k, \quad \left(M\xi = \sum_{k=1}^{\infty} x_k p_k \right).$$

б) ξ – неперервна випадкова величина.

Розіб'ємо вісь x точками xk на малі інтервали з довжиною Δxk. Зробимо серію з N незалежних випробувань і позначимо nk – число влучень випадкової величини ξ в інтервал (xk-1,xk].

$$\bar{\xi}_N \approx \sum_{k=-\infty}^{\infty} x_k \frac{n_k}{N}$$ – середнє арифметичне результатів серії випробувань.

$$\frac{n_k}{N} \approx f_\xi(x_k)\Delta x_k,$$

$$\bar{\xi}_N \approx \sum_{k=-\infty}^{\infty} x_k f_\xi(x_k)\Delta x_k \approx \int_{-\infty}^{\infty} x f_\xi(x)dx$$

Визначення. Середнім значенням (математичним сподіванням) Mξ неперервної випадкової величини ξ називається число

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

. Зауваження.

$$\int_{-\infty}^{\infty} x f_\xi(x)dx = \sum_{k=1}^{\infty} p_k \int_{-\infty}^{\infty} x\delta(x - x_k)dx = \sum_{k=1}^{\infty} p_k x_z.$$

Середнє значення функції випадкової величини

η=α(ξ).

Нехай у результаті серії з N незалежних випробувань випадкова величина ξ попадає nk разів в інтервал (xk-1, xk].

$$\bar{\eta}_N \approx \sum_{k=-\infty}^{\infty} \alpha(x_k)\frac{n_k}{N} \approx \sum_{k=-\infty}^{\infty} \alpha(x_k) f_\xi(x_k)\Delta x_k \approx \int_{-\infty}^{\infty} \alpha(x) f_\xi(x)dx,$$

$$M\eta = M\alpha(\xi) = \int_{-\infty}^{\infty} \alpha(x) f_\xi(x)dx, \qquad M\alpha(\xi) = \sum_{k=1}^{n} \alpha(x_k)p_k.$$

Нехай y = $a(x)$ монотонно зростає.

$$M\eta = \int_{-\infty}^{\infty} y f_\eta(y)dy = \int_{-\infty}^{\infty} \alpha(x) f_\xi(x) x'(y) y'(x)dx = \int_{-\infty}^{\infty} \alpha(x) f_\xi(x)dx.$$

Наприклад, η=ξ2; $f_{\xi}(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$.

Mξ2= $\int\limits_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx = \left| \begin{matrix} x=u; & dx=du \\ dv = xe^{-\frac{x^2}{2\sigma^2}}dx; & v=-\sigma^2 e^{-\frac{-x^2}{2\sigma^2}} \end{matrix} \right| = \frac{\sigma^2}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}}dx=\sigma^2$.

Середні значення функції випадкового вектора

ζ=α(ξ,η).

Розіб'ємо площину x, y на малі області Δ Gk із площами Δσк, k=1,2…, θk∈ΔGk при влученні випадкової точки (ξ,η) в область ΔGk

$\zeta = \alpha(\theta_k)$.

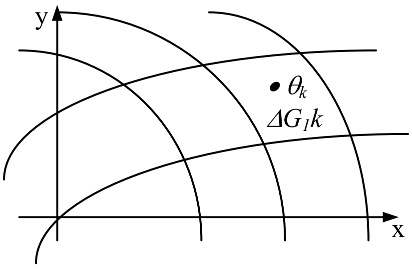
Отже,

$$\xi_N \approx \sum_{k=1}^{\infty}\alpha(\theta_k)\frac{n_k}{N} \approx \sum_{k=1}^{\infty}\alpha(\theta_k)f_{\xi,\eta}(\theta_k)\Delta\sigma_k \approx \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\alpha(x,y)f_{\xi,\eta}(x,y)d\sigma.$$

nk – число влучень випадкової точки (ξ,η) в область ΔGk.

$$M\zeta = M\alpha(\xi,\eta) = \int\int \alpha(x,y)f_{\xi,\eta}(x,y)dxdy,$$

$$M\alpha(\vec{\xi}) = \int\limits_{-\infty}^{\infty}...\int\limits_{-\infty}^{\infty}\alpha(\vec{x})f_{\vec{\xi}}(x_1,x_2,...x_n)dx_1dx_2...dx_n.$$



Властивості математичного сподівання

1. Mc = c; c – const; p(ξ = c) = 1. 2. M(kξ+b) = k

M(kξ+b) = $\int\limits_{-\infty}^{\infty}(kx+b)f_{\xi}(x)dx = kM\xi + b.$

3. M(ξ+η) = Mξ+Mη.

$$M(\xi+\eta) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x+y)f_{\xi,\eta}(x,y)dxdy = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}xf_{\xi,\eta}(x,y)dxdy + \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}yf_{\xi,\eta}(x,y)dxdy =$$

$$= \int\limits_{-\infty}^{\infty}xf_{\xi}(x)dx + \int\limits_{-\infty}^{\infty}yf_{\eta}(y)dy = M\xi + M\eta.$$

$$M(\xi_1+\xi_2+...+\xi_n)=\sum_{k=1}^{n}M\xi_k.$$

4. Для незалежних випадкових величин ξ і η

Mξη = Mξ·Mη,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_{\xi,\eta}(x,y)dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_{\xi}(x)f_{\eta}(y)dxdy=\int_{-\infty}^{\infty}xf_{\xi}(x)dx\cdot\int_{-\infty}^{\infty}yf_{\eta}(y)dy=M\xi\cdot M\eta.$$

5. p(ξ = η) = 1⇒Mξ = Mη.

p(ξ-η = 0) = 1; M(ξ-η) = 0·1 = 0 ⇒Mξ = Mη.

Приклади.

Геометричний закон розподілу

Pk=(1-p)k-1p=qk-1p; Mξ= $\frac{1}{p}$; Mξ= $\sum_{k=1}^{\infty}kq^{k-1}p=p\sum_{k=1}^{\infty}\left(q^k\right)'=p\left(\frac{1}{1-q}\right)'=p\frac{1}{(1-q)^2}=\frac{1}{p}.$

Біноміальний закон розподілу

$$p(\xi_n=k)=C_n^k p^k q^{n-k};\qquad M\xi=np.$$

Розглянемо випадкову величину ηk-число появ події А в випробуваннях з номером k (k =1,2,…n). Ця величина набуває значення 1, якщо в випробуванні з номером k подія А з'явилася, і 0 – у противному випадку.

ηk	0	1
p	q	P

$$\xi_n=\sum_{k=1}^{n}\eta_k;\qquad M\xi_n=\sum_{k=1}^{n}M\eta_k=np.$$

Закон Пуассона

p(ξ=k)= $\frac{\lambda^k}{k!}e^{-\lambda};\quad k=0,1,2,....;\qquad M\xi=\lambda;$

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

Показниковий розподіл

$$f_\tau(t)=\begin{cases}\lambda e^{-\lambda t}, t\ge 0\\ 0, \quad t<0.\end{cases}$$

$$M\tau=\int_{0}^{\infty}\lambda te^{-\lambda t}dt=\frac{1}{\lambda}.\qquad M\tau=\frac{1}{\lambda}.$$

Нормальний розподіл

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

$$M\xi=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} xe^{\frac{-(x-a)^2}{2\sigma^2}}dx=\begin{vmatrix}x-a=t\\ dx=dt\end{vmatrix}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(a+t)e^{-\frac{t^2}{2\sigma^2}}dt=a.$$

Якщо випадкова величина має симетричний розподіл щодо деякої прямої x = x0, то Mξ = x0.

2.8 Центрована випадкова величина. Дисперсія. Нормована випадкова величина

ξ0 = ξ - Mξ – центрована випадкова величина; Mξ0 = 0.

(kξ)0 = kξ0;   (ξ + k)0 = ξ0; (ξ + η)0 = ξ0 + η0.

Дисперсія випадкової величини

Дисперсія характеризує розсіювання випадкової величини щодо свого Mξ.

Величина M(ξ - Mξ) = Mξ0 = 0 не може служити характеристикою, тому що знищує додатні і від’ємні відхилення.

Можна взяти M(ξ0)n для різних n.

Якщо n = 2, маємо дисперсію:

$$D\xi = M(\xi^0)^2 = M(\xi - M\xi)^2,$$

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

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

$$D\xi \ge 0.$$

Якщо Dξ = 0, ξ - Mξ = 0 ⇒ p(ξ - Mξ = 0] = 1.

Величина ξ з імовірністю 1 набуває тільке одне значення.

$\sigma_\xi = \sqrt{D\xi}$ – середнє квадратичне відхилення випадкової величини ξ.

σ_ξ може служити мірою ширини кривої $f_\xi(x)$: чим більше σ_ξ, тим ширше крива.

Властивості дисперсії

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

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

2. $D\xi \geq 0, \quad D\xi = 0 \Rightarrow \quad \xi = M\xi = const.$

3. $D(k\xi + b) = k^2 D\xi,\ D(k\xi + b) = M\left[(k\xi + b)^0\right]^2 = Mk^2\left(\xi^0\right)^2 = k^2 M(\xi^0)^2 = k^2 D\xi.$

4. $D(\xi + \eta) = M[\xi + \eta - M(\xi + \eta)]^2 = M(\xi^0 + \eta^0)^2 = M(\xi^0)^2 + M(\eta^0)^2 + 2M\xi^0\eta^0$. Для незалежних випадкових подій:

1. D(ξ + η) = Dξ + Dη.

2. Dξ = 0, ξ = Mξ = const, M(ξ - Mξ)2 = 0, ξ = Mξ.

Приклади.

1. Геометричний розподіл.

$$p_k = (1-p)^{k-1} p = q^{k-1} p; \qquad D\xi \frac{q}{p^2};$$

$$D\xi = \sum_{k=1}^{\infty} (k - \frac{1}{p})^2 q^{k-1} p = \sum_{k=1}^{\infty} k^2 q^{k-1} p - \frac{2}{p} \sum_{k=1}^{\infty} kq^{k-1} p + \frac{1}{p^2} =$$

$$= \sum_{k=1}^{\infty} k^2 q^{k-1} p - \frac{1}{p^2} = p\sum_{k=1}^{\infty}\left(kq^k\right)' - \frac{1}{p^2} = \left(\sum_{k=1}^{\infty} kq^k\right)' p - \frac{1}{p^2};$$

$$S(q) = \sum_{k=1}^{\infty} kq^k = q + 2q^2 + 3q^3 + \ldots = q(1 + 2q + 3q^2 + \ldots) =$$

$$= q(q + q^2 + q^3 + \ldots)' = q\left(\frac{q}{1-q}\right)' = \frac{q}{(1-q)^2}.$$

$$S'(q) = \frac{(1-q)^2 + 2q(1-q)}{(1-q)^4} = \frac{p^2 + 2pq}{p^4}.$$

$$D\xi = pS'(q) - \frac{1}{p^2} = \frac{1}{p} + \frac{2q}{p^2} - \frac{1}{p^2} = \frac{p + 2q - 1}{p^2} = \frac{q}{p^2}.$$

2. Біноміальний розподіл.

$$p(\xi_n = k) = \sum_{k=1}^{n} C_n^k p^k q^{n-k}; \qquad D\xi = nqp.$$

ηk	0	1
P	q	P

ηk – число появ події А в k-тому випробуванні.

$$\xi_n = \sum_{k=1}^{n} \eta_k ;$$

$$D\eta_k = M\eta_k^2 - (M\eta_k)^2 = p - p^2 = pq;$$

$$D\xi_n = \sum_{k=1}^{n} D\eta_k = npq.$$

3. Розподіл Пуассона.

$P_k = \frac{\lambda^k}{k!} e^{-\lambda};$ k = 0,1,2,…Dξ = λ

$$D\xi = \lim_{\substack{n\to\infty \\ p\to 0}} D\xi_n = \lim_{\substack{n\to\infty \\ np\to\lambda}} np(1-p) = \lambda \lim_{n\to\infty} (1-\frac{\lambda}{n}) = \lambda$$

4. Експоненціальний закон розподілу.

$$f_\tau(t) = \begin{cases} \lambda e^{-\lambda t}, t \ge 0 \\ 0, \quad t < 0 \end{cases} \qquad D\xi = \frac{1}{\lambda^2}$$

$$D\tau = M\tau^2 - (M\tau)^2 = \lambda \int_0^\infty t^2 e^{-\lambda t} dt - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

$t^2 \div \frac{\Gamma(3)}{p^3};$ p = λ

5. Нормальний розподіл.

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

Dξ=σ2

$$D\xi = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left| \begin{matrix} x - a = u; \quad v = -\sigma^2 e^{-\frac{(x-a)^2}{2\sigma^2}} \\ (x-a) e^{-\frac{(x-a)^2}{2\sigma^2}} dx = dv; du = dx \end{matrix} \right| = \sigma^2 \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma^2$$

6. Нормована випадкова величина

$$\xi_H = \frac{\xi - M\xi}{\sigma_\xi}.$$

ξн; DξН = 1; MξН= 0.

З фізичної точки зору ξН є безрозмірною випадковою величиною.

2.9 Кореляційний момент. Коефіцієнт кореляції.
Комплексна випадкова величина. Характеристичні функції

Середнє значення ($M\xi$,$M\eta$) і дисперсія ($D\xi$, $D\eta$) випадкового вектора (ξ,η) описують властивості кожної з його проекцій окремо. При описі випадкового вектора потрібні характеристики для опису зв'язків між ξ і η.

Кореляційним моментом (коваріацією) випадкової величини ξ і η називається

$$K_{\xi,\eta} = M(\xi^0 \cdot \eta^0).$$

Властивості:

1. $K_{\xi,\eta} = K_{\eta,\xi}$.
2. $K_{\xi,\xi} = D\xi$.
3.

$$K_{\xi,\eta} = M(\xi \cdot \eta) - M\xi \cdot M\eta.$$

$$K_{\xi,\eta} = M[(\xi - M\xi)(\eta - M\eta)] = M(\xi\eta - \xi M\eta - \eta M\xi + M\xi M\eta) = M(\xi \cdot \eta) - M\xi \cdot M\eta.$$

Коефіцієнтом кореляції називається кореляційний момент нормованої випадкової величини

$$r_{\xi,\eta} = K_{\xi_H \eta_H} = M\left[\left(\frac{\xi - M\xi}{\sigma_\xi}\right)\left(\frac{\eta - M\eta}{\sigma_\eta}\right)\right] = \frac{K_{\xi,\eta}}{\sigma_\xi \sigma_\eta}.$$

Теорема.

Для будь-яких випадкових величин ξ, η коефіцієнт кореляції

$$\left|r_{\xi,\eta}\right| \le 1,$$

причому знак рівності можливий тоді і тільки тоді, коли ξ і η з ймовірністю 1 зв'язані лінійно.

$$D(\xi + \lambda\eta) = D\xi + 2\lambda M(\xi^0 \cdot \eta^0) + \lambda^2 D\eta \ge 0.$$

Квадратний тричлен відносно λ не змінює знак тоді і тільки тоді, коли

$$K_{\xi,\eta}^2 - D\xi \cdot D\eta \le 0,$$

отже, $\dfrac{K_{\xi,\eta}^2}{D\xi \cdot D\eta} \le 1$, $\left|\dfrac{K_{\xi,\eta}}{\sigma_\xi \sigma_\eta}\right| \le 1;$ тобто $\left|r_{\xi,\eta}\right| \le 1.$

Необхідність.

$$\left|r_{\xi,\eta}\right| = 1 \Rightarrow D(\xi + \lambda\eta) = 0 \Rightarrow \xi + \lambda\eta = const;$$

$$K_{\xi,\eta} = \pm\sqrt{D\xi D\eta}; \qquad \left(D\xi \pm \lambda\sqrt{D\eta}\right)^2 = 0; \qquad \lambda = \pm\frac{D\xi}{D\eta}.$$

Достатність.

$$\eta = \lambda\xi + b;$$

$$r_{\xi,\eta} = \frac{K_{\xi,\eta}}{\sigma_\xi\sigma_\eta};$$

$$K_{\xi,\eta} = M\left[\xi^0(\lambda\xi + b)^0\right] = M\xi^0 \cdot \lambda\xi^0 = \lambda D\xi;$$

$$D\eta = \lambda^2 D\xi;$$

$$r_{\xi,\eta} = \frac{\lambda D\xi}{|\lambda|\sqrt{D\xi}\sqrt{D\xi}} = \pm 1.$$

Визначення.

Випадкові величини ξ,η називаються некорельованими, якщо їх коваріація дорівнює 0. Якщо випадкові величини ξ, η незалежні, то вони некорельовані.

$$K_{\xi,\eta} = M\xi^0 \cdot M\eta^0 = 0.$$

Обернене твердження, взагалі говорячи, невірно.

Наприклад,

$$\eta = \xi^4 - \frac{1}{5}; \qquad f_\xi(x) = \begin{cases} \dfrac{1}{2} & при \quad |x| \le 1; \\ 0 & при \quad |x| > 1. \end{cases}$$

$$M\xi = \int_{-1}^{1}\frac{1}{2}xdx = 0; \qquad M\eta = \int_{-1}^{1}\frac{1}{2}(x^4 - \frac{1}{5})dx = \frac{1}{2}\left(\frac{x^5}{5} - \frac{x}{5}\right)\Bigg|_0^1 = 0,$$

$$K_{\xi,\eta} = M\xi^0\eta^0 = \int_{-1}^{1}\frac{1}{2}x\left(x^4 - \frac{1}{5}\right)dx = 0.$$

Для опису зв'язків, що існують між проекціями випадкового вектора (ξ,η) крім коваріації Kξ,η, можна використовувати числові характеристики умовних законів розподілу fξ(x| η=y), fη(y| ξ=y).

Визначення.

Умовним середнім значенням M(ξ| η=y) і умовною дисперсією D((ξ| η=y) випадкової величини ξ за умови η = y називаються вирази:

$$M(\xi|\eta = y) = \int_{-\infty}^{\infty} x f_\xi(x|\eta = y)dx;$$

$$D(\xi|\eta = y) = \int_{-\infty}^{\infty} [x - M(\xi|\eta = y)]^2 f_\xi(x|\eta = y)dx.$$

Аналогічно визначаються M(η|ξ=x) і D(η|ξ=x).

Моменти випадкової величини

nk =Mξk – початковий момент (момент) порядку k.

M|ξ|k –k-ий абсолютний момент випадкової величини.

mk = M[(ξ0)k] – центральний момент k-го порядку.

В окремому випадку $n_1 = M\xi, n_2 = M\xi^2, D\xi = n_2 - n_1^2;$

$$m_1 = M(\xi - M\xi) = M\xi - M\xi = 0, m_2 = M(\xi - M\xi)^2 = D\xi.$$

Для опису випадкового вектора також вводять початкові і центральні моменти:

nk,l = M(ξkηl); mk,l = M[(ξ0)k(η0)l].

Комплексна випадкова величина

ζ=ξ+iη є іншим способом опису випадкового вектора.

Випадкові величини ζ1 = ξ1 + iη1 і ζ2 = ξ2 + iη2 називаються незалежними, якщо є незалежними випадкові вектори (ξ1,η1) і (ξ2,η2).

$$f_{\xi_1,\eta_1}(x_1, y_1|\xi_2 = x_2, \eta_2 = y_2) = f_{\xi_1,\eta_1}(x_1, y_1);$$
$$f_{\xi_2,\eta_2}(x_2, y_2|\xi_1 = x_1, \eta_1 = y_1) = f_{\xi_2,\eta_2}(x_2, y_2);$$

Mζ=Mξ+iMη;

Dξ=M(|ζ–Mζ|2)=M(|ζ0|2);

$$K_{\zeta_1,\zeta_2} = M[(\zeta_1^0)(\bar{\zeta}_2^0)];$$

Dζ=M(|ζ|2)-|Mζ|2; Dζ=Dξ+Dη; D(aζ+b)=|a|2Dζ;

$K_{\zeta_2,\zeta_1} = \bar{K}_{\zeta_1\zeta_2}$; M(|ζ|2)≥|Mζ|, “ $\bar{\ }$ ” – операція спряження.

Характеристичні функції

Характеристичною функцією випадкової величини ξ називається середнє значення виразу eivξ.

$\varphi_\xi(v) = Me^{iv\xi}$.

$\varphi_\xi(v)$ називають також характеристичною функцією відповідного закону розподілу.

$$\varphi_\xi(v)=\begin{cases}\int\limits_{-\infty}^{\infty} e^{iv\xi} f_\xi(x)dx,\\ \sum\limits_{k=1}^{n} e^{iv\xi} p_k.\end{cases} \qquad (2.9.1)$$

Як видно з (2.9.1) характеристична функція $\varphi_\xi(v)$ є перетворенням Фур'є відповідної їй щільності ймовірності

$$f_\xi(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\varphi_\xi(v)e^{-ivx_k}dv.$$

Властивість 1

При додаванні незалежних випадкових величин їхні характеристичні функції перемножуються.

$$\varphi_{\xi+\eta}(v)=\varphi_\xi(v)\varphi_\eta(v);$$

$$\varphi_{\xi+\eta}(v)=Me^{iv(\xi+\eta)}=M(e^{iv\xi}e^{iv\eta})=Me^{iv\xi}Me^{iv\eta}=\varphi_\xi(v)\varphi_\eta(v).$$

Властивість 2

Розкладання характеристичної функції в ряд по ступенях ν дозволяє знайти всі моменти n1, n2,…випадкової величини ξ.

$$e^{iv\xi}=\sum_{k=0}^{\infty}\frac{(iv)^k}{k!}x^k,$$

$$\varphi_\xi(v)=\int\limits_{-\infty}^{\infty}\sum_{k=0}^{\infty}\frac{(iv)^k}{k!}f_\xi(x)x^k dx=\sum_{k=0}^{\infty}\frac{(iv)^k}{k!}n_k.$$

Задачі

2.9.1. Функція розподілу випадкової величини ξ має вид

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

Знайти a, b, $f\xi(x)$; $P(-1/2<\xi<1/2)$.

Рішення.

$F\xi(-1-0)=F\xi(-1+0)$; $F\xi(1-0)=F\xi(1+0)$;

$$\begin{cases}a-\frac{\pi}{2}b=0\\ a+\frac{\pi}{2}b=1; a=\frac{1}{2}; b=\frac{1}{\pi}.\end{cases} \qquad f_\xi(x)=\begin{cases}0, x<-1\\ \frac{1}{\pi\sqrt{1-x^2}}, & |x|<-1,\\ 0, & x>1;\end{cases}$$

$$P(|\xi|<\frac{1}{2})=F\left(\frac{1}{2}\right)-F\left(-\frac{1}{2}\right)=\frac{1}{2}+\frac{1}{\pi}\arcsin\frac{1}{2}=\frac{1}{2}+\frac{1}{6}=\frac{2}{3}.$$

2.9.2. Щільність ймовірності випадкової величини ξ має вид

$$f_\xi(x)=\begin{cases}0, x\le 1,\\ x-a,\ \ 1<x\le 2,\\ 0,\ \ x>2.\end{cases}$$

Знайти a, Fξ(x); Mξ, Dξ.

Рішення.

$$\int_{-\infty}^{\infty} f_\xi(x)dx=1;\quad \int_1^2 (x-a)dx=\left(\frac{x^2}{2}-ax\right)_1^2=2-2a-\frac{1}{2}+a=1;\ a=\frac{1}{2},$$

$$F_\xi(x)=\begin{cases}0,\ x\le 1,\\ \int_{-\infty}^{0} 0dx+\int_0^x (x-\frac{1}{2})dx=\dfrac{(x-\frac{1}{2})^2}{2}-\dfrac{1}{8};\ \ 0<x\le 2,\\ \int_{-\infty}^{0} 0dx+\int_0^x (x-\frac{1}{2})dx+\int_2^x 0dx=1;\ \ x>2;\end{cases}$$

Mξ= $\int_{-\infty}^{\infty} xf_\xi(x)dx=\int_1^2 x\left(x-\frac{1}{2}\right)dx=\left(\frac{x^3}{3}-\frac{x^2}{4}\right)\Bigg|_1^2=\frac{8}{3}-1-\frac{1}{3}+\frac{1}{4}=\frac{19}{12};$

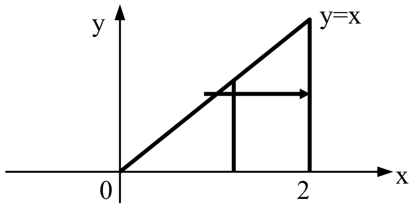
Dξ=Mξ2- (Mξ)2;

Mξ2= $\int_1^2 x^2\left(x-\frac{1}{2}\right)dx=\left(\frac{x^4}{4}-\frac{x^3}{6}\right)\Bigg|_1^2=4-\frac{4}{3}-\frac{1}{4}+\frac{1}{6}=\frac{48-16-3+2}{12}=\frac{31}{12}$;

Dξ= $\frac{31}{12}-\frac{19}{12}=\frac{372-361}{144}=\frac{11}{144}.$

2.9.3. Система двох випадкових величин (ξ, η) рівномірно розподілена в трикутнику, що обмежений прямими y=x, y=0, x=2. Знайти коефіцієнт кореляції ξ і η.

Рішення.



$$f_{\xi,\eta}(x,y)=\begin{cases}\frac{1}{2},(x,y)\in\Delta,\\ 0,(x,y)\bar{\in}\Delta.\end{cases}\qquad r_{\xi,\eta}=\frac{k_{\xi,\eta}}{\sigma_\xi\sigma_\eta}.$$

$$f_\xi(x)=\int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy=\int_0^x\frac{1}{2}dy=\frac{x}{2;}\qquad f_\xi(x)=\begin{cases}\frac{x}{2};0\le x\le 2,\\ 0,x<0;x>2.\end{cases}$$

$$f_\eta(y)=\begin{cases}\frac{y}{2};0\le y\le 2,\\ 0,y<0;y>2.\end{cases}\quad M\xi=\int_{-\infty}^{\infty}xf_\xi(x)dx=\int_0^2\frac{x^2}{2}dx=\left.\frac{x^3}{6}\right|_0^2=\frac{4}{3};\quad M\eta=\frac{4}{3};$$

$$K_{\xi,\eta}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\frac{4}{3})(y-\frac{4}{3})f_{\xi,\eta}(x,y)dxdy=\frac{1}{2}\int_0^2(x-\frac{4}{3})dx\int_0^x(y-\frac{4}{3})dy=$$

$$=\frac{1}{2}\int_0^2\left[\frac{(x-\frac{4}{3})^3}{2}-\frac{16}{9}(x-\frac{4}{3})\right]dx=\frac{1}{2}\left[\frac{\left(x-\frac{4}{3}\right)^4}{8}-\frac{8}{9}\left(x-\frac{4}{3}\right)^2\right]\Bigg|_0^2=\frac{1}{2}(\frac{2}{3^4}-\frac{32}{3^4}-\frac{32}{3^4}+\frac{128}{3^4})=\frac{11}{27}.$$

$$r_{\xi,\eta}=\frac{\frac{11}{27}}{\frac{4}{3}}=\frac{11}{9\cdot 4}=\frac{11}{36}.$$

Задано двовимірну випадкову величину (X,Y):

Y \ X	2	5	8
0,4	0,15	0,30	0,35
0,8	0,005	0,12	0,03

Знайти M(Y/X = 5).

Рішення.

Умовний закон розподілу випадкової величини Y, якщо X = x2 , дорівнює

P(Y/X=x2)= $\frac{p(Y,X=x_2)}{p(X=x_2)}=\frac{p(Y=y_j,X=x_2)}{p(X=x_2)}=\frac{p(Y=y_j,X=5)}{p(X=5)},\qquad j=1,2.$

Розподіл випадкової величини X: p(X=xi)= $\sum_{j=1}^{2}p_{ij};$

$$p(X=5)=\sum_{j=1}^{2}p(X=5;Y=y_j)=0{,}30+0{,}12=0{,}42.$$

Відповідно,

P(Y=0,4/X=5)= $\frac{0{,}30}{0{,}42}=\frac{5}{7}.$

P(Y=0,8/X=5)= $\frac{0{,}12}{0{,}42}=\frac{2}{7}.$

M(Y/X=5)= $0{,}4\cdot\frac{5}{7}+0{,}8\cdot\frac{2}{7}=\frac{2}{5}\cdot\frac{5}{7}+\frac{4}{5}\cdot\frac{2}{7}=\frac{18}{35}.$

2.10 Види збіжності. Граничні теореми теорії ймовірностей

При виконанні експерименту на його результат впливає велика кількість незалежних чи слабко залежних випадкових впливів, що приводять до випадкового характеру експерименту.
У зв'язку з цим виникає задача вивчення властивостей суми великого числа незалежних випадкових величин. Ці властивості відбиваються в граничних теоремах теорії ймовірності.

Види збіжності

Визначення. Послідовність випадкових величин ξ1, ξ2…називається збіжною до випадкової величини ξ у середньоквадратичному змісті (СК збіжність), якщо

$$\lim_{n\to\infty} M(|\xi_n-\xi|^2)=0 .$$

ξ називається СК границею {ξn}.

$$\xi = CK \lim_{n\to\infty} \xi_n , \quad \text{чи} \quad \xi = \lim_{n\to\infty} \xi_n .$$

Оскільки

$$M(|\xi_n-\xi|^2)=D(\xi_n-\xi)+(M(|\xi_n-\xi|))^2 \text{, то}$$

СК збіжність рівносильна виконанню умов:

$$\lim_{n\to\infty} M\xi_n = M\xi; \qquad \lim_{n\to\infty} D(\xi_n-\xi)=0.$$

Визначення. Послідовність випадкових величин ξ1, ξ2 ,…,ξn називається збіжною до випадкової величини ξ за ймовірностями, якщо для кожного ε>0

$$\lim_{n\to\infty} p\{|\xi_n-\xi|>\varepsilon\}=0, \qquad \text{або}$$
$$\lim_{n\to\infty} p\{|\xi_n-\xi|\le\varepsilon\}=1; \qquad p\lim_{n\to\infty}\xi_n=\xi.$$

Теорема. Для будь-якої випадкової величини ξ при кожному ε>0

$$p\{|\xi|>\varepsilon\}\le\frac{M(|\xi|^2)}{\varepsilon^2} .$$

$$p(|\xi|\ge\varepsilon)=\int_{\varepsilon}^{\infty} f_{|\xi|}(x)dx \le \int_{\varepsilon}^{\infty}\frac{x^2}{\varepsilon^2} f_{|\xi|}(x)dx \le \int_{0}^{\infty}\frac{x^2}{\varepsilon^2} f_{|\xi|}(x)dx = M\left(\frac{|\xi|^2}{\varepsilon^2}\right)=\frac{M|\xi|^2}{\varepsilon^2}.$$

$$P(|\xi|>\varepsilon)\le\frac{M|\xi|^2}{\varepsilon^2}.$$

Наслідок.

$$P(|\xi_n-\xi|>\varepsilon)\le\frac{M|\xi_n-\xi|^2}{\varepsilon^2}\xrightarrow[n\to\infty]{}0,\forall\varepsilon.$$

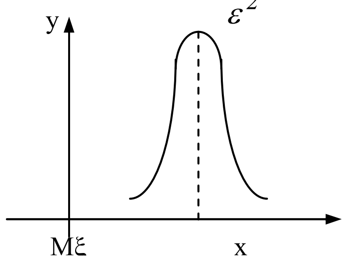
Зі збіжності в СК випливає збіжність по ймовірностей.

Нерівність Чебишова

$$P(|\xi_n-\xi|>\varepsilon)\le\frac{M|\varepsilon-M\xi|^2}{\varepsilon^2}=\frac{D\xi}{\varepsilon^2},$$

$$P(|\xi-M\xi|>\varepsilon)\le\frac{D\xi}{\varepsilon^2},$$

$$P(|\xi-M\xi|\le\varepsilon)>1-\frac{D\xi}{\varepsilon^2}. \qquad (2.10.1)$$



Як випливає з нерівностей (2.10.1), зі зменшенням дисперсії Dξ основна частина площі під кривою fξ(x) виявляється зосередженою в околі точки x = Mξ.

Зауваження. У силі загальності нерівність Чебишова дає дуже грубу оцінку ймовірностей, що входить в неї .

Наприклад,

$$P\{|\xi-M\xi|>3\sigma_\xi\}\le\frac{\sigma_\varepsilon^2}{9\sigma_\xi^2}=\frac{1}{9}=0{,}11$$

$$P\{|\xi-M\xi|>3\sigma_\xi\}=1-2\Phi(3)\approx0{,}0028$$, якщо $\xi\sim N(a,\sigma^2)$.

Визначення. Послідовність функцій розподілу F1(x),…Fn(x),.... називається збіжною до функції розподілу F(x), якщо

$$\lim_{n\to\infty}F_n(x)=F(x)$$ в усіх точках неперервності

Якщо $P\lim\limits_{n\to\infty}\xi_n=\xi$, то $\lim\limits_{n\to\infty}F_n(x)=F(x)$.

Практичне використання теорії ймовірностей засновано на наступному принципі: випадкову подію, ймовірність якої досить близька до 1, вважаємо достовірною і неможливою за дуже малої ймовірності.

Теореми, що забезпечують виконання такої схеми обробки даних, називаються законами великих чисел.

Теорема Чебишова

Нехай η1, η2…–послідовність попарно незалежних випадкових величин, дисперсії яких обмежені

$D\eta_k \le C$, $k = 1,2$ …

Тоді для кожного $\varepsilon > 0$

$$\lim_{n\to\infty} p\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\eta_k - \frac{1}{n}\sum_{k=1}^{n}M\eta_k\right| \le \varepsilon\right\} = 1.$$

$$\lim_{n\to\infty}\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\eta_n - \frac{1}{n}\sum_{k=1}^{n}M\eta_n\right| \le \varepsilon\right\} = \lim_{n\to\infty} p\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\eta_k - M\left(\frac{1}{n}\sum_{k=1}^{n}\eta_k\right)\right| \le \varepsilon\right\} \ge$$

$$\ge 1 - \frac{D\left(\frac{1}{n}\sum_{k=1}^{n}\eta_k\right)}{\varepsilon^2} \ge 1 - \frac{nc}{n^2\varepsilon^2} = 1 - \frac{c}{n\varepsilon} \underset{n\to\infty}{\to} 1.$$

Теорема Бернуллі

Нехай ξn – число появ деякої події А в серії з n незалежних випробувань, p – ймовірність появи А в окремому випробуванні.

Тоді

$$p\lim_{n\to\infty}\frac{1}{n}\xi_n = p,$$

тобто для кожного $\varepsilon>0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\xi_n - p\right| < \varepsilon\right\} = 1.$$

$$\xi_n = \sum_{k=1}^{n}\eta_k; \qquad D\eta_k = pq - обмежені; \qquad M\eta_k = p;$$

$$\frac{\xi_n}{n} = \frac{1}{n}\sum_{k=1}^{n}\eta_k; \qquad \frac{1}{n}\sum_{k=1}^{n}M\eta_k = \frac{np}{n} = p.$$

Застосовуючи теорему Чебишова, дістаємо формулу, яку шукали.

$\hat{p} = \frac{\xi n}{n} \rightarrow$ p при необмеженому числі випробувань. Збіг теоретичних випробувань із закономірностями, що фактично спостерігаються, свідчить про правильну схему побудови теорії ймовірності.

Центральна гранична теорема

Нехай ξ1,ξ2,…послідовність незалежних випадкових величин з дисперсією D1, D2,…Dn…Треті абсолютні центральні моменти їх обмежені mk=M|ξk - Mξk|3 ≤ C

Тоді випадкова величина

$$S_n = \frac{\xi_1 + \xi_2 + ...\xi_n - [\, M\xi_1 + M\xi_2 + ... + M\xi_n}{\sqrt{D_1 + D_2 + ... + D_n}}$$

розподілена асимптотично нормально із середнім Mξn = 0 і DSn = 1, тобто

P(α <Sn <β) → Φ(β) - Φ(α) при n → ∞.

Теорема Муавра-Лапласа (інтегральна гранична теорема)

Нехай ξn – число появ деякої події A в серії з n незалежних випробувань, p – ймовірність появи події A в окремому випробуванні. Тоді

$$\lim_{n\to\infty} P\left\{\alpha \le \frac{\xi_n - np}{\sqrt{np(1-p)}} < \beta\right\} = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{t^2}{2}} dt.$$

Теорема дозволяє при досить великих n визначити

$$P(X1 \le \xi \le X2) = \sum C_n^k p^k q^{n-k} = P\left(\frac{x_1 - np}{\sqrt{npq}} < \frac{\xi_n - np}{\sqrt{npq}} < \frac{x_2 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{x_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{x_1 - np}{\sqrt{npq}}\right).$$

Приклад. Обчислити P(725 < ξn < 715) число появ герба в 1500 киданнях.

$$p = \frac{1}{2};\ np = 1500 \cdot \frac{1}{2} = 750; \quad npq = 375.$$

$$P(725 < \xi_{1500} < 715) = \Phi\left(\frac{25}{\sqrt{375}}\right) - \Phi\left(\frac{-25}{\sqrt{375}}\right) = 2\Phi\left(\frac{25}{\sqrt{375}}\right) \approx 0{,}803.$$

Задачі

2.10.1. Дискретна випадкова величина X задана законом розподілу

X	0,3	0,6
p	0,2	0,8

Використовуючи нерівності Чебишова, оцінити ймовірність того, що $|X-MX|<0,2$.

ηдо	0	1
p(ηк)	q	p

Рішення: Скористаємося нерівністю Чебишова у формі

p(|X - MX| $\le$ ε) $\geq 1-\frac{DX}{\varepsilon^2}$.

Знайдемо MX і DX.

MX = 0,3·0,2 + 0,6·0,8 = 0,54;

DX = MX2 - M2X = (0,32·0,2 + 0,62·0,8 - 0,542 = 0,0144.

ε = 0,2;

p(|X - 0,54| $\le$ 0,2) $\geq 1-\frac{0,0144}{0,04}=0,64$.

2.10.2. Пристрій складається з 10 елементів, що працюють незалежно. Імовірність відмови кожного елемента за час Т дорівнює 0,05. За допомогою нерівності Чебишова оцінити імовірність того, що абсолютна величина різниці між числом елементів, що відмовили, і середнім числом відмов за час Т не менше двох.

Рішення. Число елементів ξ_{10}, що відмовили, має біноміальний розподіл:

P(ξ10 = к) = $C_{10}^k p^k q^{n-k}$; p = 0,05; q = 0,95.

За нерівностістю Чебишова p(|ξn - Mξn| > ε) $\le$ $\frac{D\xi_n}{\varepsilon^2}$.

У нашому випадку Mξn = np = 10·0,05 = 0,5;

Dξn = npq = 10·0,05·0,95 = 0,475.

Отже, p(|ξ10 - 0,5|> 2) $\le$ $\frac{0,475}{4}=0,88$.

2.10.3. Імовірність народження хлопчика 0,51. Знайти ймовірність того, що серед 100 немовлят виявиться 50 хлопчиків.

Рішення. Число новонароджених хлопчиків ξn має біноміальний розподіл, але

ξn = $\sum_{k=1}^{n}\eta_k$, де ηk=0, якщо немовля – дівчинка і ηk=1, якщо немовля – хлопчик.

На підставі центральної граничної теореми при n >> 1 ξn має нормальний розподіл, для якого $M\xi n = np = 100\cdot 0{,}51 = 51$,

$D\xi n = npq = 100\cdot 0{,}51\cdot 0{,}49 \approx 25$,

тому

$$p(\xi 100=50)=\frac{1}{\sqrt{2\pi D\xi_n}}e^{-\frac{(x-M\xi_n)^2}{2D\xi_n}}=\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-51)^2}{50}}\approx 0{,}046.$$

2.10.4. Імовірність появи події в кожному з незалежних випробувань дорівнює 0,8. Скільки потрібно зробити випробувань, щоб з ймовірністю 0,9 можна було чекати, що подія з'явиться більше 75 разів?

Рішення. За умови p = 0,8; q = 0,2;

$$p(75 < \xi_n < \infty) = \Phi(\infty) - \Phi\left(\frac{75 - np}{\sqrt{npq}}\right) = 0{,}9.$$

Оскільки $\Phi(\infty) \approx 0{,}5$, то

$$0{,}9 = 0{,}5 - \Phi\left(\frac{75-0{,}8n}{0{,}4\sqrt{n}}\right) \Rightarrow \Phi\left(\frac{75-0{,}8n}{0{,}4\sqrt{n}}\right) = -0{,}4; -0.4 = \Phi(1{,}28).$$

Дістаємо рівняння

$$\left(\frac{75-0{,}8n}{0{,}4\sqrt{n}}\right) = -1{,}28,$$

розв’язавши яке відносно $\sqrt{n}$, дістаємо $\sqrt{n}$ = 10, звідки n = 100.

3 МАТЕМАТИЧНА СТАТИСТИКА
3.1 ВИБІРКОВИЙ МЕТОД

При вивченні якісної чи кількісної ознаки деякого об'єкта (наприклад, стандартність деталі – якісна ознака, контрольований розмір деталі – кількісна) проводиться серія n незалежних вимірів, зв'язаних з даною ознакою $(x_1^1, x_2^1, \ldots, x_n^1)$. При повторних n незалежних випробуваннях одержимо нові виміри $(x_1^2, x_2^2, \ldots, x_n^2)$ і т.д. Результат вимірів пов'язаний з наявністю сумарного числа слабко залежних і незалежних впливів, тому кожен результат виміру є випадковим. Можна говорити про деякий n–вимірний випадковий вектор $(X_1,\dots,X_n)$ вимірів, причому кожен компонент цього вектора характеризує ту саму випадкову величину X.

Генеральною сукупністю називається ймовірносний простір $(\Omega(\omega), B, P)$ (тобто простір елементарних подій із заданим на ньому полем подій B і ймовірностями P) із визначеною на цьому просторі випадковою величиною X. Одиницею генеральної сукупності називається елементарна подія і значення випадкової величини, що відповідає їй.

Випадковою вибіркою чи просто вибіркою об'єму n називається послідовність $X_1,\dots,X_n$ незалежних однаково розподілених випадкових величин, розподіл кожної з яких збігається з розподілом випадкової величини X, що досліджується. Конкретний набір вибіркових значень $x_1, x_2,\dots,x_n$ варто розглядати як реалізацію n-мірного випадкового вектора $(X_1,\dots,X_n)$ з функцією розподілу:

$$F(x_1,\dots,x_n) = P(X_1 < x_1,\dots,X_n < x_n) = \prod_{i=1}^{n} F_i(x_i).$$

Метод вивчення генеральної сукупності на основі вибірки називається вибірковим. Оскільки результати вимірів включають систематичну і випадкові помилки; при одержанні достовірної інформації про досліджувану ознаку ці помилки повинні бути зведені до мінімуму.

Нехай a – ознака, що досліджується, а $\hat{a}(n)$ – її оцінка, що отримана вибірковим методом. Точність оцінки буде характеризувати величина $M(\hat{a}(n)-a)^2$ (M – математичне сподівання), у якій виділимо систематичну і випадкові помилки.

$$M(\hat{a}(n)-a)^2 = M\left[(\overset{A}{a}(n) - M\hat{a}(n)) + (M\hat{a}(n) - a)\right]^2 =$$

$$= M(\hat{a}(n) - M\hat{a}(n))^2 + M^2(M\hat{a}(n) - a)^2.$$

Перший доданок в останньому виразі характеризує випадкову помилку, другий - систематичну. Для усунення систематичної помилки зажадаємо, щоб $M\hat{a}(n) = a$. Умова ($M(\hat{a}(n) - M\hat{a}(n))^2 \underset{n\to\infty}{\to} 0$ забезпечить зменшення випадкової помилки при збільшенні об'єму n вибірки.

У статистиці умову $M\hat{a}(n) = a$ називають умовою незсуненості оцінки, умову $D\hat{a}(n) = M\left(\hat{a}(n) - M\hat{a}(n)\right)^2 \underset{n\to\infty}{\to} 0$ – спроможністю оцінки. До числа статистичних властивостей оцінок відносять ефективність оцінки, за якою оцінка має мінімальну дисперсію у визначеному класі оцінок.

Застосуємо вибірковий метод для одержання оцінок математичного сподівання і дисперсії випадкової величини X.

Нехай $X_1, \ldots, X_n$ - випадкова вибірка. Як вибіркову оцінку $MX = a$ математичного сподівання X розглянемо величину

$$\hat{a} = \frac{X_1 + \ldots + X_n}{n}. \qquad (3.1.1)$$

Перевіримо, чи задовольняє ця оцінка вимогам незсуненості і спроможності.

$$M\hat{a} = M\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n}\left(MX_1 + \ldots + MX_n\right) = \left|\frac{1}{n}\sum_{i=1}^{n} MX_i = a\right| = a;$$

$$D\hat{a} = D\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n^2}\left(DX_1 + \ldots + DX_n\right) = \left|DX_i = \sigma^2\right| = \frac{\sigma^2}{n} \underset{n\to\infty}{\to} 0.$$

Зручно ввести оператор усереднення $\hat{M} = \frac{1}{n}\sum_{i=1}^{n}(\cdot), \hat{M}X = \frac{1}{n}\sum_{i=1}^{n} X_i$. У цьому випадку

$$MX \approx \hat{M}X = \hat{a}. \qquad (3.1.2)$$

Для одержання вибіркової оцінки $DX = M(X - MX)^2$ замінимо оператор математичного сподівання оператором усереднення. Тоді

$$\hat{\sigma}^2 = \hat{M}\left(X - \hat{M}X\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \hat{a}\right)^2, \qquad (3.1.3)$$

де $\hat{a} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Застосуємо до $\hat{\sigma}^2$ оцінки операцію математичного сподівання, використовуючи в (3.1.3) тотожнє перетворення

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

$$= M\left\{\frac{1}{n}\sum_{i=1}^{n}\left[(X_i-a)^2-2(X_i-a)(\hat{a}-a)+(\hat{a}-a)^2\right]\right\} =$$

$$= \frac{1}{n}\sum_{i=1}^{n}M(X_i-a)^2-\frac{1}{n}\sum_{i=1}^{n}M(\hat{a}-a)^2 = \sigma^2-\frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2 .$$

Зсунення в оцінці (3.1.3) можна ліквідувати перетворенням

$$M\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \sigma^2 .$$

У статистиці реалізації незсунених оцінок математичного сподівання і дисперсії позначають відповідно $\overline{x}$ і s2, а самі оцінки $\overline{x}$ і s^2. Вони виходять з (3.1.1) і (3.1.3), за відповідного нормування, якщо замість випадкової вибірки підставити конкретну вибірку $(x_1,\ldots,x_n)$, тобто $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$, і називаються відповідно емпіричним (вибірковим) середнім і емпіричною (вибірковою) дисперсією.

Аналогічно знайдемо вибіркову оцінку коваріації випадкових величин X і Y - $k_{XY} = M\left[(X-MX)(Y-MY)\right]$. Нехай $(X_1,\ldots,X_n)$ і $(Y_1,\ldots,Y_n)$ - вибірки з випадкових величин X і Y, тоді

$$\hat{k}_{XY} = \hat{M}(X-\hat{M}X)(Y-\hat{M}Y) = \hat{M}\hat{X}^0\hat{Y}^0 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y}) ,$$

де $\overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i$; $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n}Y_i$.

Вибіркові оцінки асиметрії й ексцесу випадкової величини відповідно дорівнюють

$$\hat{a}_s = \frac{\hat{M}\left(\hat{X}^0\right)^3}{\hat{\sigma}_x^2} ; \quad \hat{e}_k = \frac{\hat{M}\left(\hat{X}^0\right)^4}{\hat{\sigma}_x^2}-3 .$$

3.1.1 МЕТОД МОМЕНТІВ

Метод моментів полягає у використанні вибіркових моментів для визначення параметрів розподілу випадкової величини X. Застосуємо, наприклад, метод моментів для визначення параметрів рівномірного закону розподілу з щільністю ймовірностей:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, x \in [a,b], \\ 0,\ x \overline{\in} [a,b]. \end{cases}$$

Обчислимо математичне сподівання і дисперсію:

$$MX = \int_a^b \frac{x}{b-a} dx = \frac{1}{2} \frac{b^2 - a^2}{b-a} = \frac{b+a}{2}, \qquad (3.1.1.1)$$

$$DX = \int_a^b \frac{x^2}{b-a} dx - \left(\frac{b+a}{2} \right)^2 = \frac{b^3 - a^3}{3(b-a)} - \frac{(b+a)^2}{4} =$$
$$= \frac{4\left(b^2 + ab + a^2\right) - 3\left(b^2 + 2ab + b^2\right)}{12} = \frac{(b-a)^2}{12}. \qquad (3.1.1.2)$$

Для визначення оцінок параметрів a і b замінимо в рівняннях (3.1.1.1) і (3.1.1.2) MX і DX їхніми оцінками $\overline{X}$ і S^2. Дістаємо систему рівнянь $\overline{X} = \dfrac{\widehat{b} + \widehat{a}}{2}$, $S = \dfrac{\widehat{b} - \widehat{a}}{2\sqrt{3}}$, звідки знаходимо:

$$\widehat{a} = \overline{X} - \sqrt{3}S; \ \ \widehat{b} = \overline{X} + \sqrt{3}S.$$

3.1.2 МЕТОД НАЙБІЛЬШОЇ ПРАВДОПОДІБНОСТІ

Нехай дана вибірка $(X_1, \ldots, X_n)$ об'єму n з генеральної сукупності з неперервно розподіленою випадковою величиною X. Нехай щільність ймовірності X містить невідомий параметр θ, який потрібно оцінити за вибіркою, і має вигляд $f(x,\theta)$.

Функцією правдоподібності називають функцію параметра θ, що визначена співвідношенням

$$L(X_1,...,X_n,\theta)=f(X_1,\theta)\cdot f(X_2,\theta)\cdot ...\cdot f(X_n,\theta). \qquad (3.1.2.1)$$

Розглянемо випадок дискретної випадкової величини X з можливими значеннями $x_1, x_2, ...$ та імовірностями $P(X=x_j)=p_j(\theta)$. Позначимо через X_r найбільше з можливих значень, що зустрічається у вибірці, а через $n_1, n_2, ..., n_r$ - абсолютні частоти, з якими з'являються значення $X_1, ..., X_r$ у вибірці $\left(\sum_{i=1}^{r} n_i = n\right)$. У цьому випадку функцією правдоподібності називають функцію параметра θ, що визначена співвідношенням:

$$L(X_1,...,X_n,\theta)=p_1^{n_1}(\theta)p_2^{n_2}(\theta)...p_r^{n_r}(\theta). \qquad (3.1.2.2.)$$

Метод найбільшої правдоподібності полягає в тому, що за оцінку параметра береться значення, при якому функція правдоподібності досягає свого максимуму. Параметр θ знаходять, розв'язуючи відносно θ рівняння:

$$\frac{\partial L}{\partial\theta}=0, \theta=\hat{\theta}. \qquad (3.1.2.3)$$

Часто замість (3.1.2.3) використовують рівняння:

$$\frac{1}{L}\frac{\partial L}{\partial\theta}=\frac{\partial \ln L}{\partial\theta}=0, \qquad \theta=\hat{\theta} \qquad (3.1.2.4)$$

Якщо щільність $f(x,\theta_1,...,\theta_e)$ чи ймовірності $p(X=x_j)=p_j(\theta_1,...,\theta_e)$ залежать від l параметрів, то найбільш правдоподібну оцінку системи параметрів $\theta_1,...,\theta_e$ одержують розв'язанням системи рівнянь:

$$\frac{\partial L}{\partial \theta_j} = 0 \; j = (1,...,l), \theta_j = \hat{\theta}_j \qquad (3.1.2.5)$$

або

$$\frac{\partial \ln L}{\partial \theta_j} = 0 \; j = (1,...,l), \theta_j = \hat{\theta}_j . \qquad (3.1.2.6)$$

Найбільш правдоподібні оцінки мають деякі цікавіі властивості. За досить загальних умов вони є спроможними й асимптотично нормально розподіленими (однак не завжди незсуненими), мають серед всіх асимптотично нормально розподілених оцінок найбільшу ефективність. Справедливо наступне положення: якщо взагалі є ефективна оцінка, то вона отримується методом найбільшої правдоподібності.

Задачі

3.1.2.1. Оцінити ймовірність p деякої події A.
Нехай

$$X = \begin{cases} 0, якщо \quad A \quad не\ здійснилося \\ 1, якщо \quad A \quad здійснилось \end{cases}$$

Рішення. $P(X = 0) = 1 - p$; $P(X = 1) = p$. Нехай в n незалежних спостереженнях подія A відбулася k раз, тобто $n_1 = k; n_0 = n - k$. Таким чином, згідно (3.1.2.2), маємо $L = p^k (1-p)^{n-k}$, $\frac{\partial \ln L}{\partial p} = \frac{k}{p} - \frac{n-k}{1-p} = 0$. Звідси випливає, що $\hat{p} = \frac{k}{n}$. Отже, $\hat{p}$ є найбільш правдоподібною оцінкою параметра p . Випадкова величина k біноміально розподілена, $M\hat{p} = M\frac{k}{n} = \frac{np}{n} = p$;

$D\hat{p} = D\frac{k}{n} = \frac{1}{n^2} Dk = \frac{pq}{n} \to 0 \; при \; n \to \infty.$ Отже, $\hat{p}$ - незсунена оцінка ймовірності, асимптотично спроможна і асимптотично нормальна.

3.1.2.2. Нехай випадкова величина X розподілена за законом Пуасона з невідомим параметром λ. Проведемо вибірку й одержимо значення $X_1,\dots,X_n$ (X_i – цілі числа). Нехай r – найбільше з чисел, що спостерігаються у вибірці, $n_0, n_1, n_2, \dots, n_r$ - абсолютні частоти, з якими числа $0,1,2,\dots,r$ з'являються у вибірці $p_j(\lambda) = p(X=j) = \frac{\lambda^j}{j!}e^{-\lambda}$. Тоді, відповідно до формули (3.1.2.2), $L = \prod_{j=0}^{r}\left(\frac{\lambda^j}{j!}e^{-\lambda}\right)^{n_j}$. Зі співвідношення (3.1.2.4) дістаємо $\frac{\partial \ln L}{\partial \lambda} = \sum_{j=0}^{r} n_j\left(\frac{j}{\lambda}-1\right) = 0$,

звідки $\hat{\lambda} = \frac{\sum_{j=0}^{r} jn_j}{\sum_{j=0}^{r} n_j} = \frac{1}{n}\sum_{j=0}^{n} X_j = \overline{X}$.

Величина $\hat{\lambda} = \overline{X}$ є, таким чином, правдоподібною оцінкою для λ і разом з тим спроможна, асимптотично нормально розподіленою.

3.1.2.3. Нехай випадкова величина X розподілена нормально з параметрами a і σ^2. Їх належно оцінювати, виходячи із вибірки $(X_1,\dots,X_n)$ об'єму n.

Рішення. Функція правдоподібності

$$L\left(X_1,\dots,X_n;a,\sigma^2\right) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(X_k-a)^2\right\},$$

отже,

$$\ln L = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(X_k-a)^2.$$

Згідно (3.1.2.6), дістаємо наступні рівняння для визначення a і σ^2:

$\frac{\partial \ln L}{\partial a} = \frac{1}{\sigma^2}\sum_{k=1}^{n}(X_k-a) = 0$; $\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{k=1}^{n}(X_k-a)^2 = 0$, звідки

$\hat{a} = \frac{1}{n}\sum_{k=1}^{n}X_k$ і $\hat{\sigma}^2 = \frac{1}{n}\sum(X_k-\hat{a})^2$. Отже, $\left(\hat{a},\hat{\sigma}^2\right)$ є найбільш правдоподібною оцінкою параметрів $\left(a,\sigma^2\right)$. Ми вже знаємо, що $\hat{\sigma}^2$ не є незсуненою оцінкою, а тільки асимптотично незсуненою.

3.2 Довірчі оцінки

Вибіркова оцінка, яка розглянута в попередньому розділі, будучи точковою, дає оцінки для значення відповідного параметра з даної вибірки, але нічого не говорить про точність і вірогідность оцінки. Такі дані надають довірчі оцінки. Нехай $\left(X_1,\ldots,X_n\right)$ випадкова вибірка з генеральної сукупності з випадковою величиною X, розподіл якої залежить від параметра γ. Нехай $\Gamma_1\left(X_1,\ldots,X_n\right),\Gamma_2\left(X_1,\ldots,X_n\right)$ – такі функції вибірок, що при довільному γ виконується рівність:

$$P\left(\Gamma_1\left(X_1,\ldots,X_n\right)<\gamma<\Gamma_2\left(X_1,\ldots,X_n\right)\right)=1-\alpha. \qquad (3.2.1)$$

Тоді випадковий інтервал $\left(\Gamma_1,\Gamma_2\right)$ називається довірчою оцінкою параметра γ з мірою надійності $1-\alpha$ (з рівнем значущості α).

Якщо є реалізація $\left(x_1,\ldots,x_n\right)$ вибірки $\left(X_1,\ldots,X_n\right)$, то реалізація довірчої оцінки дає довірчий інтервал $\left(\gamma_1,\gamma_2\right)$ і у великому ряді вибірок істинне значення лежить приблизно у $(1-\alpha)100\%$ випадків усередині обчислених довірчих границь γ_1 і γ_2. Рівність (3.2.1) можна інтерпретувати і так: випадковий інтервал $\left(\Gamma_1,\Gamma_2\right)$ "покриває" істинний параметр із довірчою ймовірністю $1-\alpha$.

У математичній статистиці часто використовують поняття квантилей, процентних точок (однобічних критичних границь і двосторонніх критичних границь). Квантиллю рівня p чи p-квантиллю x_p випадкової величини X з функцією розподілу $F(x)$ називається розв'язок рівняння $F(x)=p$.

Однобічною критичною границею, що відповідає рівню значущості α (процентною точкою рівня $q=\alpha\cdot 100\%$) неперервної випадкової величини X з функцією розподілу $F(x)$, називається значення випадкової величини x_α, для якої $P\left(X>x_\alpha\right)=\alpha$, чи $F\left(x_\alpha\right)=1-\alpha$.

Нижньою і верхньою критичними границями, що відповідають рівню значущості α неперервної випадкової величини X з функцією розподілу $F(x)$, називаються значення випадкової величини $x_{1-\alpha/2}$ і $x_{\alpha/2}$, для яких $P\left(X > x_{1-\alpha/2}\right) = 1 - \frac{\alpha}{2}$;

$P\left(X > x_{\alpha/2}\right) = \frac{\alpha}{2}$;

$$P\left(x_{1-\frac{\alpha}{2}} < X < x_{\frac{\alpha}{2}}\right) = F\left(x_{\frac{\alpha}{2}}\right) - F\left(x_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

Для симетричних випадкових величин, у яких щільності розподілу симетричні щодо деякої точки a, нижні і верхні критичні границі задовольняють умови $x_{1-\alpha/2} + x_{\alpha/2} = 2\alpha$, що дає можливість приводити таблиці лише для процентних точок чи квантилей, більших ніж a. Так, для стандартної нормальної випадкової величини з рівнем значущості $\alpha = 0{,}05$ $x_{0,975} = -x_{0,025}$.

Квантиль, однобічні і двосторонні критичні границі зображені на рис.3.2.1.

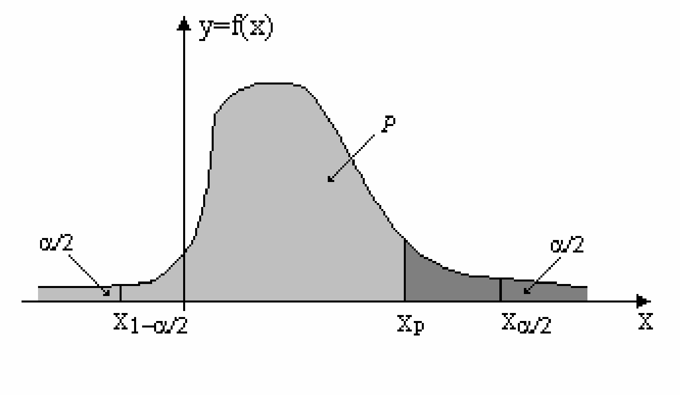


Рис. 3.2.1 p-квантиль і критичні точки для закону розподілу $f(x)$.

3.2.1. Довірча оцінка невідомої ймовірності за великими вибірками

Частота $\hat{p} = \hat{p}(A)$ є точковою оцінкою $p = p(A)$, вона асимптотично нормально розподілена з $M\hat{p} = p$ і $D\hat{p} = D\left(\frac{k}{n}\right) = \frac{npq}{n^2} = \frac{pq}{n}$.

Якщо $X \in N(x; a, \sigma)$, то $\dfrac{X-a}{\sigma} \in N(x;0,1)$. Задамо α. Величина $x_{\alpha/2}$ така, що

$$P\left(-x_{\alpha/2} < \frac{X-a}{\sigma} < x_{\alpha/2}\right) = \frac{1}{\sqrt{2\pi}} \int_{-x_{\alpha/2}}^{x_{\alpha/2}} e^{-\frac{x^2}{2}} dx = 2\Phi\left(x_{\alpha/2}\right) = 1-\alpha$$

може бути знайдена з рівняння $2\Phi(x_{\alpha/2}) = 1-\alpha$ за допомогою таблиць для функцій Лапласа.

Це ж міркування застосуємо до $\hat{p}$. За заданим α можна знайти $x_{\alpha/2}$ так, щоб

$$\Phi\left(\frac{\left|\hat{p}(A) - p\right|}{\sqrt{pq/n}} < x_{\alpha/2}\right) \approx 1-\alpha$$

. З нерівності $\dfrac{\left|\hat{p}(A) - p\right|}{\sqrt{pq/n}} < x_{\alpha/2}$ випливає, що

$$p^2\left(1 + \frac{x_{\alpha/2}}{n}\right) - p\left(2\hat{p} + \frac{x_{\alpha/2}^2}{n}\right) + \hat{p}^2 < 0$$

, звідки можна обчислити два значення $p_1(\hat{p}(A))$ і $p_2(\hat{p}(A))$, що представляють довірчі оцінки для p. Якщо α обрано досить малим, то випадковий інтервал $(p_1(\hat{p}(A)), p_2(\hat{p}(A)))$ "покриває" p майже напевно.

3.2.2 Довірчі оцінки для параметрів нормального закону

3.2.2.1 Довірча оцінка a при відомому σ

$$\hat{a} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
, $X_i \in N(x; a, \sigma^2)$, тоді $\hat{a} \in N(x; a, \dfrac{\sigma^2}{n})$.

Відповідно,

$$\frac{(\hat{a} - a)\sqrt{n}}{\sigma} \in N(x;0,1)$$
.

Для стандартної нормальної випадкової величини з рівнем значущості α нижня і верхня критичні границі відповідно дорівнюють $-x_{\frac{\alpha}{2}}$ і $x_{\frac{\alpha}{2}}$.

Маємо

$$P(-x_{\frac{\alpha}{2}} < \frac{(\tilde{a}-a)\sqrt{n}}{\sigma} < x_{\frac{\alpha}{2}}) = 2\Phi(x_{\frac{\alpha}{2}}) = 1-\alpha,$$

або

$$P(\tilde{a} - \frac{\sigma x_{\frac{\alpha}{2}}}{\sqrt{n}} < a < \tilde{a} + \frac{\sigma x_{\frac{\alpha}{2}}}{\sqrt{n}}) = 1-\alpha$$

.

$$f(x) = \frac{1}{\sqrt{2n}} e^{\frac{-x^2}{2}}$$

.

Таким чином, $(\tilde{a} - \frac{\sigma x_{\alpha/2}}{\sqrt{n}}, \tilde{a} + \frac{\sigma x_{\alpha/2}}{\sqrt{n}})$ - довірча оцінка для параметра а з мірою надійності $1-\alpha$.

Приклад 1.Знайти довірчий інтервал з надійністю 0,95 для оцінки невідомого математичного сподівання а нормально розподіленої ознаки X генеральної сукупності, якщо дано: генеральне середнє квадратичне відхилення $\sigma = 5$, вибіркова середня $\overline{x} = 14$ і об'єм вибірки n=25.

Рішення. Знайдемо $x_{\frac{\alpha}{2}} = x_{0,025}$ з $2\Phi(x_{0,025}) = 0,95$ чи співвідношення $\Phi(x_{0,025}) = 0,475$, звідки $x_{0,025} = 1,96$.

Підставимо в довірчу оцінку значення $\tilde{a} = \overline{x} = 14; \sigma = 5; n = 25$, дістаємо довірчий інтервал

$12,04 < a < 15,96$.

3.2.2.2 Довірча оцінка a при невідомому σ

Оцінка заснована на тому факті, що при висловлених припущеннях величина

$\frac{\tilde{a}-a}{\tilde{\sigma}}\sqrt{n}$ задовольняє t- розподілу з n-1 ступенями свободи.

Визначаючи однобічну критичну точку $t_{\frac{\alpha}{2}} = t_{\frac{\alpha}{2}}(n-1)$ з умови

$$P\left(\frac{|\hat{a}-a|\sqrt{n}}{s} > t_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$$

, одержимо довірчу оцінку для а у виді

$$\hat{a} - \frac{st_{\frac{\alpha}{2}}}{\sqrt{n}} < a < \hat{a} + \frac{st_{\frac{\alpha}{2}}}{\sqrt{n}}.$$

Для конкретної вибірки $(x_1,...,x_n)$ об'єму n довірча оцінки для а є її довірчим інтервалом.

3.2.2.3 Довірча оцінка σ при невідомому a

Відправною точкою є той факт, що при заданих передумовах величина $\frac{(n-1)\hat{\sigma}^2}{\sigma^2}$ задовольняє χ^2 - розподілу з n-1 ступенями свободи. За заданим рівнем значущості α і $k = n-1$ ступенями свободи знаходимо критичні точки χ_1^2 і χ_2^2 розподілу χ^2 такі, що

$$P(\chi^2 < \chi_1^2) = \frac{\alpha}{2}, \quad P(\chi^2 > \chi_2^2) = \frac{\alpha}{2},$$

$$P(\chi_1^2 < \frac{(n-1)\hat{\sigma}^2}{\sigma^2} < \chi_2^2) = 1-\alpha$$

, чи

$$P\left(\frac{(n-1)\hat{\sigma}^2}{\chi_2^2} < \sigma^2 < \frac{(n-1)\hat{\sigma}^2}{\chi_1^2}\right) = 1-\alpha.$$

Таким чином, $\left(\frac{(n-1)\hat{\sigma}^2}{\chi_2^2}, \frac{(n-1)\hat{\sigma}^2}{\chi_1^2}\right)$ є довірчою оцінкою σ^2 з мірою надійності $1-\alpha$.

Приклад 2. Зроблено 12 вимірів деякої фізичної величини приладом (без систематичної помилки), причому $\hat{\sigma}$=s=0.36 - виправлене середнє квадратичне відхилення випадкових помилок вимірів. Знайти довірчий інтервал, що покриває σ з заданою ймовірністю $\alpha = 0,99$ і k=11.

Рішення. Знайдемо критичні точки розподілу:

χ^2 : $\chi_1^2 = \chi^2(0{,}995;11) = 3{,}1; \chi_2^2 = \chi^2(0{,}005;11) = 26{,}8$.

Довірчі границі відповідно дорівнюють:

$$\sqrt{\frac{(n-1)s^2}{\chi_2^2}} = \sqrt{\frac{11\cdot 0{,}36}{26{,}8}} = 0{,}38; \sqrt{\frac{(n-1)s}{\chi_1^2}} = \sqrt{\frac{11\cdot 0{,}36}{3{,}1}} = 1{,}13 .$$

Отже, $0{,}38 < \sigma < 1{,}13$.

3.3 Перевірка статистичних гіпотез

Задача перевірки гіпотез виникає в тих випадках, коли поставлене питання вимагає відповіді «так » чи «ні». Вибір одної з двох відповідей відбувається в умовах невизначеності, тому будь-який вибір може виявитися помилковим. Задача полягає в тому, щоб вибрати відповіді в якомусь смислі найкращі.

Нехай H0 - висунута гіпотеза щодо параметра закону чи функції розподілу. H0 називають нульовою чи основною гіпотезою, H1 -конкуруючою чи альтернативною гіпотезою.

У підсумку застосування гіпотез можуть бути допущені помилки двох типів. Помилка першого роду - відкинута правильна гіпотеза. Імовірність помилки першого роду називають рівнем значущості і позначається вона α. Помилка другого роду - буде прийнята хибна гіпотеза. Імовірність помилки другого роду позначимо β (рис.3.3.1). Як випливає з рис 3.3.1, одночасна мінімізація α і β неможливі. Величина 1-α називається надійністю критерію, а величина 1-β- потужністю критерію (ймовірність відкидання основної гіпотези, якщо вірна конкуруюча).

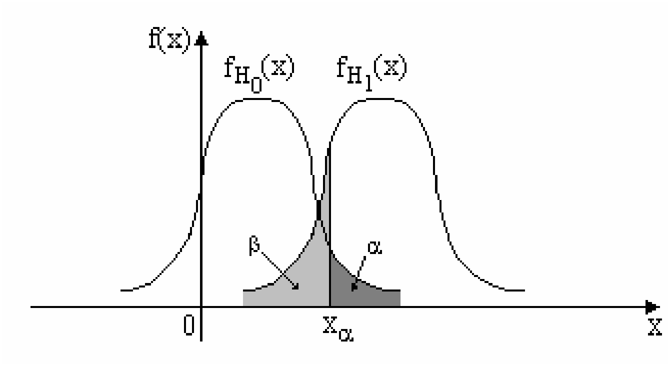


Рис.3.3.1 - Ймовірності помилок $\alpha = \int_{x_\alpha}^{\infty} f_{H_0}(x)dx$ $\quad$ Ц $\beta = \int_{-\infty}^{x_\alpha} f_{H_1}(x)dx$

Для побудови оптимального критерію виходять з того, що помилки першого і другого роду нерівноправні за своїми наслідками.

Критерій рівня значущості α називається оптимальним, якщо він забезпечує мінімум ймовірності помилки другого роду (максимум потужності) серед усіх критеріїв того ж рівня значущості.

Гіпотеза називається простою, якщо їй відповідає один розподіл чи одна точка простору параметрів, і складною, якщо вона зводиться до вибору розподілу з їх множини або точки з інтервалу (скінченого чи нескінченного).

Оптимальний критерій не завжди існує. Він існує завжди, якщо гіпотеза проста.

Існує зв'язок між критерієм для перевірки гіпотез і довірчими інтервалами . У випадку перевірки простої гіпотези H0 : $\Theta = \Theta_0$ проти альтернативи $\Theta \neq \Theta_0$ довірчий інтервал, заснований на оцінці найбільшої правдоподібності для Θ, дає найкращий критерій перевірки гіпотези H0 проти H1.

Нехай $\hat{\Theta} = \hat{\Theta}(X_1,...,X_n)$ - оцінка найбільшої правдоподібності (о.н.п.) параметра $\Theta = \Theta_0$, отримана на основі вибірки $(X_1,...,X_n)$ з випадкової величини X.

$$1-\alpha = P(\Gamma_1(\hat{\Theta}) < \Theta < \Gamma_2(\hat{\Theta})) = P(c_1 < K(X_1,...,X_n) < c_2).$$

Значення c1 і c2 є критичними для закону розподілу випадкової величини $K(X_1,...,X_n)$, що залежить тільки від вибірки . Її розглядають як статистичний критерій. Нехай $K(x_1,...,x_n)$ - значення критерію, що спостерігається, a - рівень значущості, k_α - критична точка. Перевірка гіпотези H0 : Θ=Θ0 проти альтернатив H1 : Θ≠Θ0 ; H2 : Θ>Θ0 ; H3 : Θ<Θ0 визначається наступними правилами :

H1 : Θ≠Θ0 , якщо $K(x_1,...,x_n) < k_{1-\alpha/2}$ і $K(x_1,...,x_n) > k_{\alpha/2}$,

H2 : Θ>Θ0 , якщо $K(x_1,...,x_n) > k_\alpha$,

H3 : Θ<Θ0 , якщо $K(x_1,...,x_n) < k_{1-\alpha}$.

У противному випадку говорять, що вибірки, що розглядаються, не суперечать гіпотезі H1 : Θ=Θ0 .

3.3.1 Порівняння вибіркової середньої з гіпотетичною генеральною середньою нормальної сукупності

A. Дисперсія генеральної сукупності відома.

Нехай випадкова величина $X \in N(x;a,\sigma^2)$, a_0 - гіпотетичне (передбачуване) значення. Висувається гіпотеза H0 : $a = a_0$ проти конкуруючої H1 : $a \neq a_0$.

Нехай $\hat{a} = \frac{1}{n}\sum_{i=1}^{n} X_i$ - о.н.п. для довільної вибірки (X1,...,Xn) з ознакою $X \in N(x;a,\sigma^2)$.

Відповідно, при a=a0

$$K = \frac{(\hat{a} - a_0)\sqrt{n}}{\sigma} \in N(k;0,1), \qquad (3.3.1.1)$$

а двосторонні критичні точки стандартного нормального розподілу дорівнюють

$$\pm k_{\frac{\alpha}{2}} \left(\int_{K_{\frac{\alpha}{2}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{\alpha}{2} \right).$$

Нехай (х1,...,хп) - конкретна вибірка з випадкової величини $X \in N(x;a,\sigma^2)$. Тоді значення критерію, що спостерігається

$$k_{спост.} = K(x_1,...,x_n) = \frac{(\tilde{a}(x_1,...,x_n) - a_o)\sqrt{n}}{\sigma}.$$

Для прийняття гіпотез одержуємо наступне вирішальне правило:

$$|k_{спост.}| \begin{cases} \le k_{\frac{\alpha}{2}} \to приймається \quad гіпотеза \quad H_0 \\ > k_{\frac{\alpha}{2}} \to приймається \quad гіпотеза \quad H_1 \end{cases} \qquad (3.3.1.2)$$

Гіпотеза Н1 : $a \ne a_0$ еквівалентна умовам $\{a < a_0 \text{ або } a > a_0\}$. Критерій $k_{спост.} < -k_{\frac{\alpha}{2}}$ відповідає прийняттю гіпотези $a < a_0$, а $k_{спост.} > k_{\frac{\alpha}{2}}$ - гіпотезі $\alpha > \alpha_0$.

Нехай перевіряється гіпотеза Н0 : $a = a_0$ проти конкуруючої Н1 : $\alpha < \alpha_0$ $(\alpha > \alpha_0)$ з рівнем значущості α. Знаходиться лівостороння (правобічна) критична точка - $k_\alpha(k_\alpha)$ випадкової величини (3.3.1.1), а вирішальне правило приймає вид :

$$k_H \begin{cases} \ge -k_\alpha(\le k_\alpha) \to приймається \quad гіпотеза \quad H_0 \\ < -k_\alpha(> k_\alpha) \to приймається \quad гіпотеза \quad H_1 \end{cases}$$

Б. Дисперсія генеральної сукупності невідома.

Як величину σ в (3.3.1.1) використовують її незсувну оцінку $\tilde{\sigma}$. Критерієм (3.3.1.1) є t- розподілена випадкова величина з n-1 ступенями свободи:

$$t(n-1) = \frac{(\tilde{a} - a_0)\sqrt{n}}{\tilde{\sigma}} \qquad (3.3.1.3)$$

У вирішальному правилі прийняття гіпотез використовуються критичні точки t(n-1) -розподілу рівня значущості α.

Приклад 1. Проектний контрольований розмір деталей, що виготовляються верстатом-автоматом, $a = a_0 = 35мм$. Виміри 20 випадково відібраних виробів дали наступні результати :

контрольований розмір	x_i	34,8	34,9	35,0	35,1	35,3
частота (число виробів)	ni	2	3	4	6	5

Потрібно при рівні значущості 0,05 перевірити нульову гіпотезу H0: $a = a_0 = 35мм$ при конкуруючій гіпотезі $H_1 : a \neq 35$.

Рішення. Знайдемо середній розмір виробів вибірки :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{5} n_i x_i = \frac{2 \cdot 34,8 + 3 \cdot 31,9 + 4 \cdot 35,0 + 6 \cdot 35,1 + 5 \cdot 35,3}{20} = 35,07$$

Знайдемо емпіричну дисперсію :

$$s^2 = \frac{1}{n}(\sum_{i=1}^{5} n_i x_i^2 - n\bar{x}^2) = 0,022; \qquad s = \sqrt{0,022} = 0,15$$

Значення критерію(3.3.1.3.), що спостерігається за вибіркою:

$$t_{спост.} = \frac{(\bar{x} - a_0)\sqrt{n}}{s} = \frac{(35,07 - 35,0)\sqrt{20}}{0,15} = 2,15$$

.

За умови критична область - двостороння; для рівня значущості $\alpha = 0,05$ $t(0,025;19) = 2,09.$

$t_{спост.} > t(0,025;19) = 2,09$ -нульову гіпотезу відкидаємо. Іншими словами, верстат не забезпечує проектного розміру виробів і вимагає налагодження.

3.3.2 Порівняння виправленої вибіркової дисперсії з гіпотетичною генеральною дисперсією нормальної сукупності

Позначимо через n обсяг вибірки, за якою знайдена емпірична дисперсія s2.

Правило 1. Для того, щоб при заданому рівні значущості α перевірити нульову гіпотезу H0: σ2=σ20 про рівність невідомої генеральної дисперсії σ2 передбачуваному значенню $\sigma20$ при конкуруючій гіпотезі H1: $\sigma2\neq\sigma20$, треба обчислити значення критерію, що спостерігається:

$$\chi^2_{спост.} = \frac{(n-1)s^2}{\sigma_0^2}$$

і за таблицею критичних точок розподілу $\chi2$ за заданим рівнем значущості α і числом ступенів свободи k=n-1 знайти ліву $\chi_1^2 = \chi^2\left(1-\frac{\alpha}{2},k\right)$ і праву

$\chi_2^2 = \chi^2\left(\frac{\alpha}{2},k\right)$ критичні точки.

Якщо $\chi12\leq\chi2спост.\leq\chi22$, немає підстав відкинути нульову гіпотезу. Якщо $\chi2спост.<\chi12$ чи $\chi2спост>\chi22$ - нульову гіпотезу відкидають.

Правило 2. При альтернативній гіпотезі H1: $\sigma2<\sigma20$ ($\sigma2>\sigma20$) знаходять критичну точку $\chi2(1-\alpha;k)$ ($\chi2(\alpha;k)$).

Якщо $\chi2спост\geq\chi2(1-\alpha;k)$ ($\chi2спост\leq\chi2(\alpha;k)$) , немає підстав відкинути нульову гіпотезу.

Якщо $\chi2спост<\chi2(1-\alpha;k)$ ($\chi2спост>\chi2(\alpha;k)$) , нульову гіпотезу відкидають.

3.3.3 Порівняння відносної частоти, що спостерігається, з гіпотетичною ймовірністю появи події

Нехай за досить великим числом n незалежних випробувань, у кожному з яких ймовірність p появи події постійна, але невідома, знайдена відносна частота $\hat{p} = \frac{m}{n}$. Потрібно при заданому рівні значущості α перевірити нульову гіпотезу, яка полягає в тому, що невідома ймовірність p дорівнює гіпотетичній ймовірності p0.

$$U_{спост} = \frac{(\hat{p}-p_0)\sqrt{n}}{\sqrt{p_0q_0}} \in N(u;0,1), як\ np_0q_0 > 9.$$

$$\Phi(U_{кp}) = \frac{1-\alpha}{2},$$

де Ф - функція Лапласа.

Для перевірки гіпотези H0 : p=p0 проти конкуруючої p≠p0 справедливо правило 1:

$$\left|U_{спост.}\right|\begin{cases}\le U_{кр} \to H_0,\\ > U_{кр} \to H_1.\end{cases}$$

При конкуруючій гіпотезі H1: p<p0 (p>p0) застосовують правило 2:

$$\Phi(U_{к\,р}) = \frac{1-2\alpha}{2};$$

$$U_{спост.}\begin{cases}\ge -U_{к\,р}\left(\le U_{ђ\,р}\right) \to H_0,\\ < -U_{к\,р}\left(> U_{ђ\,р}\right) \to H_1.\end{cases}$$

Приклад 2. Партія виробів приймається, якщо ймовірність того, що виріб виявиться бракованим, не перевищує 0,02. Серед випадково відібраних 480 виробів виявилося 12 дефектних. Чи можна прийняти партію? Рівень значущості прийняти α=0,05.

Рішення. Нульова гіпотеза H0 має вид:

H0: p=p0=0,02; конкуруюча H1: p>0,02.

$$\hat{p} = \frac{m}{n} = \frac{12}{480} = 0{,}025$$

$$U_{спост} = \frac{(\hat{p} - p_0)\sqrt{n}}{\sqrt{p_0 q_0}} = \frac{(0.025 - 0.02)\sqrt{400}}{\sqrt{0.02 \cdot 0.98}} = 0.71$$

$$\Phi\left(U_{к\,р}\right) = \frac{1 - 2\cdot 0{,}05}{2} = 0.45 \Rightarrow U_{к\,р} = 1.645$$

$$U_{спост} < U_{к\,р} \to H_0.$$

Таким чином, партію можна прийняти.

3.3.4 Перевірка гіпотези про значущість вибіркового коефіцієнта кореляції

Нехай двовимірна генеральна сукупність (X,Y) розподілена нормально. З цієї сукупності витягнута конкретна вибірка об'єму n : (x1,y1),...,(xn,yn), і за її допомогою знайдемо вибірковий коефіцієнт кореляції €. Потрібно перевірити нульову гіпотезу H0: r=0 про рівність нулю генерального коефіцієнта кореляції , тобто некорельованість випадкових величин X і Y.

Обчислимо значення випадкової величини Стьюдента, що спостерігається, з n-2 ступенями свободи:

$$T_{спост} = \frac{€\sqrt{n-2}}{\sqrt{1-€^2}}, де\ € = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{[\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2]^{1/2}};$$

$$\bar{x} = \frac{1}{n}\sum x_i;\ \bar{y} = \frac{1}{n}\sum y_i.$$

Нехай альтернативна гіпотеза H1: r≠0; критична точка t(α/2;n-2) для двосторонньої критичної області.

Вирішальне правило:
$$\left|T_{спост}\right| \begin{cases} \le t\left(\frac{\alpha}{2}, n-2\right) \to H_0, \\ > t\left(\frac{\alpha}{2}, n-2\right) \to H_1. \end{cases}$$

3.3.5 Перевірка гіпотез про рівність середніх і дисперсій двох генеральних сукупностей

Розглядаються критерії про рівність середніх і дисперсій двох генеральних сукупностей, що мають нормальні розподіли. Вибірку з першої генеральної сукупності будемо позначати X=(X1,..Xn), а з другої Y=(Y1,..Ym).

3.3.5.1 Гіпотеза про рівність дисперсій при невідомих середніх

Розглянемо гіпотезу H0: σ2X=σ2Y проти гіпотези H1: σ2X≠σ2Y. Сформульовану задачу можна замінити їй еквівалентною:

$$\left\{H_0 : \frac{\sigma_X^2}{\sigma_Y^2} = 1; H_1 : \frac{\sigma_X^2}{\sigma_Y^2} \ne 1;\right\}.$$

Тепер гіпотеза H0 проста. Оскільки середні невідомі, то найкращими оцінками дисперсій σX2 і σY2 є:

$$€_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\ Ц\ €_Y^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \bar{Y})^2.$$

Розглянемо статистику:

$$F = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} \; (\hat{\sigma}_X^2 > \hat{\sigma}_Y^2). \qquad (3.3.5.1.1)$$

При відношенні дисперсій (σX2/σY2)=1 відношення їхніх оцінок повинно бути близьким до 1, тому критерій виглядає наступним чином:

- гіпотеза H0 відкидається, якщо

$$\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} < c_1 \text{ або } \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} > c_2 \quad (c_1 < 1 < c_2).$$

Залишається знайти константи c1 і c2. Для цього потрібно знати розподіл (3.3.5.1.1). Цей розподіл відомий і табульований. Він називається розподілом Фішера з n1=n-1 і n2=m-1 ступенями свободи. Для критичних точок рівня значущості α справедливе співвідношення F(α;n1,n2)=1/F(1-α;n2,n1), звідки

$$c_1 = F(1-\frac{\alpha}{2};n_1,n_2) = 1 \Big/ F(\frac{\alpha}{2};n_2,n_1)$$

$$c_2 = F(\frac{\alpha}{2};n_1,n_2).$$

У випадку альтернативи σX2>σY2 розглядається однобічна критична границя F-статистики.

3.3.5.2 Гіпотеза про рівність дисперсій при відомих середніх

У даному випадку

$$\hat{\sigma}_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a_X)^2 \text{ Џ } \hat{\sigma}_Y^2 = \frac{1}{m}\sum_{i=1}^{m}(Y_i - a_Y)^2,$$

де ax і ay -відомі середні генеральних сукупностей. Якщо вірна гіпотеза H0: σX2=σY2=σ2 , то $\frac{\hat{\sigma}_X^2}{\sigma^2}$ і $\frac{\hat{\sigma}_Y^2}{\sigma^2}$ розподілені за законом χ2 , відповідно, з n і m ступенями свободи. Тому статистика $\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}$ розподілена за законом Фішера з n і m ступенями свободи.

3.3.5.3 Гіпотеза про рівність середніх при невідомих дисперсіях

Припускається, що σ2x=σ2y=σ2, причому як величина дисперсії σ2, так і середні ах і ау невідомі. Потрібно перевірити основну гіпотезу Н0: ах=ау при конкуруючій Н1: ах ≠ ау.

Гіпотеза Н0 справедлива при будь-яких а=ах і а=ау, тобто вона складна, але може бути зведена до простої, якщо розглядати різницю середніх ах-ау. Різниця оцінок âx –ây розподілена нормально, тому що нормальні самі оцінки âx і ây:

$$D(\hat a_X-\hat a_Y)=\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}=\sigma^2\,\frac{n+m}{nm},$$

$$\frac{\hat a_X-\hat a_Y}{\sqrt{D(X-Y)}}=\frac{\hat a_X-\hat a_Y}{\sigma}\sqrt{\frac{nm}{n+m}}\in N(z;0.1)$$

Оцінимо невідому величину σ. Скористаємося тим, що

$$\hat\sigma_X^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\hat a_X)^2,\text{Ъ}\hat\sigma_Y^2=\frac{1}{m-1}\sum_{i=1}^{m}(Y_i-\hat a_Y)^2.$$

Відомо, що

$$(n-1)\frac{\hat\sigma_X^2}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\hat a_X^2}{\sigma}\right)^2\ \text{Ц}\ (m-1)\frac{\hat\sigma_Y^2}{\sigma^2}=\sum_{i=1}^{m}\left(\frac{Y_i-\hat a_Y^2}{\sigma}\right)^2$$

розподілені за законом χ2 з n-1 і m-1 ступенями свободи,відповідно, тому

$$(n-1)\frac{\hat\sigma_X^2}{\sigma^2}+(m-1)\frac{\hat\sigma_Y^2}{\sigma^2}$$

розподілена за законом χ2 з n+m-2 ступенями свободи.

Розглянемо

$$\frac{(n-1)\hat\sigma_X^2+(m-1)\hat\sigma_Y^2}{n+m-2}=\frac{(n-1)\hat\sigma_X^2}{n+m-2}+\frac{m-1}{n+m-2}\hat\sigma_Y^2.$$

Оскільки $M\hat\sigma_X^2=M\hat\sigma_Y^2=\sigma^2$, то

$$M\left[\frac{(n-1)\hat\sigma_X^2}{n+m-2}+\frac{(m-1)\hat\sigma_Y^2}{n+m-2}\right]=\frac{(n-1)\sigma^2}{n+m-2}+\frac{(m-1)\sigma^2}{n+m-2}=\sigma^2,$$

отже,

$$\hat\sigma^2=\frac{n-1}{n+m-2}\hat\sigma_X^2+\frac{m-1}{n+m-2}\hat\sigma_Y^2,$$

є незсувною оцінкою для σ.

Тепер одержуємо, що величина

$$t_{спост} = \frac{\tilde{a}_X - \tilde{a}_Y}{\sqrt{\left[(n-1)\tilde{\sigma}_X^2 + (m-1)\tilde{\sigma}_Y^2\right] / (n+m-2)}}\sqrt{\frac{nm}{n+m}} \qquad (3.3.5.3.1.)$$

залежить тільки від вибірки і розподілена за законом Стьюдента з n+m-2 ступенями свободи.

З викладеного можна одержати наступний критерій рівня значущості α для перевірки гіпотези H0: ax=ay проти альтернативи H1: ax≠ay. Гіпотеза H0 відкидається, якщо |tспост|>tдв(α;n+m-2), де tдв(α;n+m-2)- критична точка, що відповідає рівню значущості для двосторонньої критичної області.

3.3.5.4 Гіпотеза про рівність середніх при відомих дисперсіях

Критерієм є випадкова величина:

$$Z_{спnст} = \frac{\tilde{a}_X - \tilde{a}_Y}{\sqrt{DX + DY}} = \frac{(\tilde{a}_X - \tilde{a}_Y)\sqrt{n}}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \in N(z;0.1)$$

.

Тому гіпотеза H0: ач =ay проти альтернативної H1: ач ≠ay відкидається при рівні значущості α, якщо |Zспост|> $z_{\alpha/2}$, де $z_{\alpha/2}$ - критична границя стандартного нормального розподілу, тобто P(|Zспост|> $z_{\alpha/2}$)=α.

Приклад 3. Із двох партій виробів, виготовлених на двох однаково побудованих верстатах, витягнуті малі вибірки, обсяги яких n=10 і m=12. Отримано наступні результати

Контрольований розмір виробів першого верстата	xi	3.4	3.5	3.7	3.9
Частота (число виробів)	ni	2	3	4	1
Контрольований розмір виробів другого верстата	yi	3.2	3.4	3.6	
Частота	mi	2	2	8	

Припускається, що випадкові величини X і Y розподілені нормально. Потрібно при рівні значущості 0.02 перевірити гіпотезу H0: ax=ay про рівність середніх розмірів виробів при конкуруючій гіпотезі ax≠ay.

Рішення. Знайдемо вибіркові середні і дисперсії:

$$\overline{x} = \frac{\sum n_i x_i}{n} = 3{,}6; \ \overline{y} = \frac{\sum m_i y_i}{m} = 3{,}5; \ s_x^2 = \frac{\sum n_i (x_i - \overline{x})^2}{n-1} = 0{,}0267;$$

$$s_y^2 = \frac{\sum m_i (y_i - \overline{y})^2}{m-1} = 0{,}0254.$$

Дисперсії різні. Порівняємо їх, використовуючи F - статистику (3.3.5.1.1)

$$F_{спост} = \frac{s_x^2}{s_y^2} = \frac{0{,}0267}{0{,}0254} = 1{,}05$$

Знайдемо критичні точки F - статистики:

$$c_1 = F_1(0{,}99;9{,}11) = \frac{1}{F(0{,}01;11{,}9)} = \frac{1}{5{,}18} = 0{,}19$$

$$c_2 = F_2(0{,}01;9{,}11) = 4{,}63; \qquad 1{,}05 \in [0{,}19;4{,}63]$$

- дисперсії розрізняються незначною мірою, тобто можна вважати σx2=σy2.

Обчислимо значення критерію Стьюдента, що спостерігається, за формулою (3.3.5.3.1):

$$t_{спост} = \frac{3{,}6 - 3{,}5}{\sqrt{(9 \cdot 0{,}0267 + 11 \cdot 0{,}0254)/20}} \sqrt{\frac{120}{22}} = 0{,}72; \ t_{Љ‰}(0{,}02;20) = 2{,}53; \ 0{,}72 < 2{,}53.$$

Таким чином, середні розміри виробів істотно не розрізняються.

3.4 Оцінка функції розподілу і щільності розподілу

Нехай X1,..,Xn - вибірка з генеральної сукупності, яка представлена випадковою величиною X. Спостереження розташовують у порядку їхнього зростання. Упорядкована послідовність X1*<X2*<...<Xn* називається варіаційним рядом, а значення Xi* - варіантами. Якщо n велике, спостереження поєднують у m груп, як правило $m = \overline{6{,}10}$, указуючи число варіантів (частот), що потрапили в кожну групу.

Статистичним розподілом вибірки називають перелік варіант варіаційного ряду (послідовності інтервалів) і відповідних їм частот ni ($\sum ni=n$), чи відносних частот $p_i = \frac{n_i}{n}$.

Емпіричною функцією розподілу (функцією розподілу вибірки) називають функцію F*(x), що визначає для кожного значення x відносну частоту події X<x;

$$F^*(x) = \frac{n_x}{n},$$

де nx - число варіант, менших за x; n – об'єм вибірки.

F*(x) задовольняє наступним властивостям:

1.0 ≤ F*(x) ≤ 1;

2. F*(x) - неспадна функція;

3. Якщо x1 - найменша варіанта, а xk - найбільша, то F*(x) = 0 при x ≤ x1 і F*(x) = 1 при x > xk.

Приклад 1. Знайти емпіричну функцію розподілу за даним розподілом вибірки

xi	1	4	6
ni	10	15	25

Рішення. Об'єм вибірки n = 10 + 15 + 25 = 50.

Найменша варіанта xi = 1, отже

$$F^*(x) = P(X < x) = \begin{cases} 0, & x \le 1; \\ \frac{10}{50} = 0{,}2, & 1 < x \le 4; \\ \frac{25}{50} = 0{,}5, & 4 < x \le 6; \\ 1, & x > 6. \end{cases}$$

Графік цієї функції відображений на рис. 3.4.1.

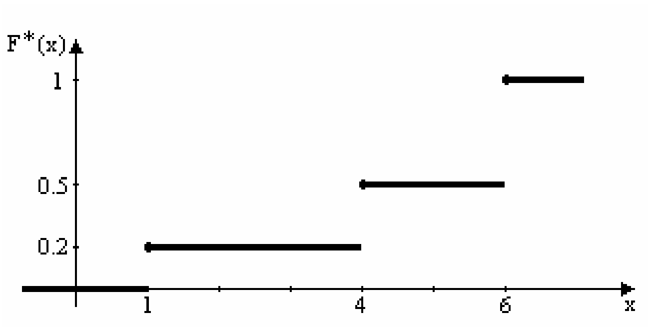


Рис.3.4.1

При неперервному розподілі ознаки весь інтервал, в якому укладені всі спостереженні значення ознаки, розбивають на ряд часткових інтервалів з довжиною hi і знаходять ni - суму частот варіант, що потрапили в кожний і-ий інтервал.

Гістограмою частот називають східчасту фігуру, що складається з прямокутників, основами яких є часткові інтервали з довжиною hi, а їх висоти дорівнюють частки ni/hi (щільність частоти). Площа часткового і-ого прямокутника дорівнює (hi ni)/hi = ni - сумі частот варіант, що потрапили в і-й інтервал. Площа гістограми частот дорівнює сумі всіх частот, тобто об'єму вибірки n.

Гістограмою відносних частот називають східчасту фігуру, що складається з прямокутників, основами яких є часткові інтервали з довжиною hi , а їх висоти дорівнюють частки $\frac{\hat{p}_i}{h_i}$ (щільність відносної частоти). Площа часткового і-го прямокутника дорівнює $h_i \cdot \frac{\hat{p}_i}{h_i} = \hat{p}_i$ - відносній частоті варіант, що потрапили в і-ий інтервал. Площа гістограми відносних частот дорівнює сумі всіх відносних частот, тобто одиниці, і представляє собою оцінку щільності розподілу неперервної випадкової величини X.

Полігон частот будують точно так само, як гістограму, але середини $\bar{x}_i$ верхніх граней прямокутників гістограми з'єднують прямими лініями.

Для випадкових величин із гладкою щільністю розподілу полігон відносних частот краще оцінює щільність.

Полігоном відносних частот для дискретної ознаки X називають ламану, що з'єднує точки $\left(\bar{x}_i, \hat{p}_i\right), \ldots, \left(\bar{x}_n, \hat{p}_n\right)$.

Приклад 1. У таблиці 1 приведені згруповані дані по запасах взуття в 100 магазинах.

Таблиця 1

Інтервал тис. грн.	0,2-2,2	2,2-4,2	4,2-6,2	6,2-8,2	8,2-10,2	10,2-12,2
Число магазинів в інтервалі ni	70	20	4	3	2	1

Знайти емпіричну щільність розподілу вартості взуття методом гістограми і полігона відносних частот.

Рішення. Гістограма і полігон відносних частот надані на рисунку 3.4.2

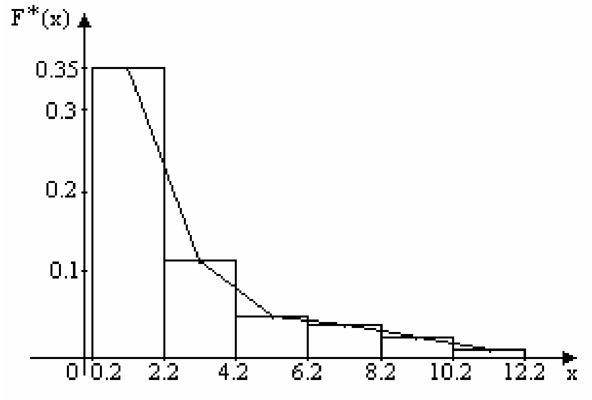


Рис. 3.4.2

На рисунку 3.4.2 вздовж осі x відкладені інтервали спостереження з довжиною hi = h = 2 , а із середин інтервалів паралельно осі у відкладені щільності відносних частот ni/2n . Суцільна лінія - гістограма, пунктирна - полігон відносних частот.

3.5 Критерії згоди

Нехай (X1,..,Xn) - вибірка з невідомим законом розподілу F(X). Розглянемо гіпотезу H0: F(x)=F0(x) при гіпотезі H1: F(x)≠F0(x), що конкурує з нею, . F0(x) - деяка задана функція розподілу.

Задача перевірки гіпотез щодо законів розподілу називається задачею перевірки згоди, а критерій для цієї задачі - критерієм згоди.

Розглянемо критерій згоди $\chi 2$ чи критерій Пірсона.

Розіб'ємо вісь x на m інтервалів $(-\infty,a_1],(a_1,a_2],(a_2,a_3],...,(a_m,\infty)$. Якщо істинна функція розподілу F(x) збігається з F0(x), то при великих n

$\frac{n_i}{n} \approx p_i$.

Розглянемо випадкову величину (ni - випадкове)

$$\chi^2_{спост.} = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i},$$

при $n \to \infty$ вона прямує до χ2 - розподілу випадкової величини з т-е-1 ступенями свободи (е- число статистичних параметрів).

Вирішальне правило для рівня значущості α:

$$\chi^2_{спост} \begin{cases} \le \chi^2_{кр} \to приймається\ г\ гіпоте\ \ H_0, \\ > \chi^2_{кр} \to відкидається\ \ \ гіпотеза\ H_0. \end{cases}$$

$$P\left(\chi_{спост}{}^2 > \chi^2_{кр}\right) = \alpha; \quad n \ge 50 \div 60; \quad m \ge 5 \div 8.$$

При побудові χ2спост. повинна виконуватись умова ni ≥ 10; у противному випадку поєднують інтервали.

У випадку застосування гіпотези H0 говорять, що розходження між F(x) і F0(x) є випадковим з довірчою ймовірністю 1-α і обумовлено обмеженістю вибірки.

Приклад 1. Проаналізувати розподіл валиків (у мм), що виготовлені на токарському верстаті. Для аналізу відібрані n = 200 валиків, результати яких згруповані в 10 інтервалів довжини h = 0,2 мм. Згруповані дані наведені в таблиці 1.

Таблиця 1

Інтервали xi-1-xi	3,1-3,3	3,3-3,5	3,5-3,7	3,7-3,9	3,9-4,1	4,1-4,3	4,3-4,5	4,5-4,7	4,7-4,9	4,9-5,1
Частоти ni	1	5	4	18	86	62	14	6	3	1

Рішення. Застосуємо гіпотезу щодо нормальності закону розподілу діаметрів валика, тобто

$$F_0(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Знайдемо значення о.н.п. параметрів а і σ2 , прийнявши в якості варіанта вибірки середини часткових інтервалів $\bar{x}_i = \frac{x_{i-1} + x_i}{2}$. Дістанемо

$$\tilde{a} = \frac{1}{n}\sum_{i=1}^{m} n_i \bar{x}_i = \frac{1}{200}(1 \cdot 3{,}2 + 5 \cdot 3{,}4 + \ldots + 1 \cdot 5{,}0) = 4{,}08.$$

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{m} n_i(\bar{x}_i - \tilde{a})^2 = \frac{1}{200}\left(1(3{,}2-4{,}08)^2 + .. + 1(5{,}0-4{,}8)^2\right) = 0{,}06.$$

$$\tilde{\sigma} = \sqrt{0{,}06} = 0{,}25$$

Таким чином,

$$F_n(x) = \frac{1}{\sqrt{2\pi}0{,}25} e^{\frac{-(x-4{,}08)^2}{2 \cdot 0{,}006^2}}.$$

Схема обчислення критерію χ2 наведена в таблиці 2

Таблиця 2

i	Інтервали	ni	i	Інтервали	ni	npi	ni-npi	$\frac{(n_i - np_i)^2}{np_i}$
1	(∞;3,3]	1	1	(-∞;3,7]	10	12,9	-2,9	0,65
2	(3,3;3,5]	5	2	(3,7;3,9]	18	34,3	-16,3	7,75
3	(3,5;3,7]	4	3	(3,9;4,1]	86	59,2	26,8	12,13
4	(3,7;3,9]	18	4	(4,1;4,3]	62	55,7	-0,63	0,71
5	(3,9;4,1]	86	5	(4,3;4,5]	14	28,6	14,6	7,45
6	(4,1;4,3]	62	6	(4,5;+∞)	10	9,3	0,7	0,05
7	(4,3;4,5]	14						
8	(4,5;4,7]	6						
9	(4,7;4,9]	3						
10	(4,9;+∞)	1						

χ2спост=28,74

Оскільки ni ≥ 10, перші й останні три інтервали об'єднані.

$$np_1 = 200\left[\Phi\left(\frac{3,7-4,08}{0,25}\right)-\Phi(-\infty)\right] = 200\left[\frac{1}{2}-\Phi(1,52)\right] = 200(0,5-0,4357) = 12,9;$$

$$np_2 = 200\left[\Phi\left(\frac{3,9-4,08}{0,25}\right)-\Phi\left(\frac{3,7-4,08}{0,25}\right)\right] = 34,3;$$

. . . .

$$np_6 = 200\left[\Phi(\infty)-\Phi\left(\frac{4,9-4,08}{0,25}\right)\right] = 200\left[0,5-\Phi\left(\frac{0,82}{0,25}\right)\right] = 0,1.$$

Оскільки при обчисленні χ2спост використані два статистичних параметри $\hat{a}$ и $\hat{\sigma}^2$, тому статистика χ2 має m-e-1= 6-2-1=3 ступені свободи. Критична точка рівня значущості a = 0,05 дорівнює χ2кр(6;0,05) = 7,8; 28,74 > 7,8 , отже, за критерієм Пірсона, гіпотеза про нормальність закону розподілу діаметрів валика відхиляється.

Критерій згоди Колмогорова можна застосовувати в тому випадку, коли параметри теоретичного закону розподілу F0(x) визначаються не за даними вибірки, що досліджується. Дослідне значення D0 величини D визначається формулою:

D=max|Fn(x)-F0(x)|,

де Fn(x) - емпірична, а F0(x) - теоретична функції розподілу; n ≥ 40 ÷ 50 .

Випадкова величина $\sqrt{n}D$ залежить від результатів спостережень і підкоряється закону Колмогорова. Знаючи критичну точку закону Колмогорова рівня значущості $\alpha,\ P\left(\sqrt{n}D > \lambda_\alpha\right) = \alpha,$ дістанемо критерій прийняття гіпотези:

$\sqrt{n}D$ ≤ λα - приймається гіпотеза H0;

$\sqrt{n}D$ > λα - відхиляється гіпотеза H0.

При порівнянні з законом розподілу F0(x) для експериментально визначених параметрів беруть рівень значущості α = 0,2 чи навіть 0,3.

Застосуємо критерій Колмогорова, використовуючи дані таблиці 1. В результаті одержимо таблицю 3.

Таблиця 3

i	ni	npi	nFn(x)	nF0(x)	n|Fn(x)-F0(x)|
1	10	12,9	10	12,9	2,9
2	18	34,3	28	47,2	19,2
3	86	59,2	114	106,4	7,6
4	62	55,7	176	162,1	13,9
5	14	28,6	190	190,7	0,7
6	10	9,3	200	200	0

nD=max n|Fn(x)-F0(x)|=19,2

Критерієм є випадкова величина:

$$\lambda = \sqrt{n}D; \quad \lambda_{спост} = \frac{19,2}{\sqrt{200}} = 1,36.$$

Приймемо рівень значущості $\alpha = 0,2$. З таблиці критичних точок розподілу Колмогорова знаходимо $\lambda(0,2) = 1,07$.

1,36 > 1,07, отже, критерій Колмогорова також підтверджує наявність істотних розбіжностей між досліджуваним і нормальним розподілом.

3.6 Парна кореляція

Кореляційна залежність двох випадкових величин задається моделлю X=X(Y,Z), Y=Y(X,Z), де Z - набір зовнішніх випадкових факторів.

Парна кореляція вивчає характеристики взаємозв'язку двох випадкових величин. Підставою одержання цих характеристик служить спільний розподіл випадкових величин:

$F(x,y)=P(X<x,Y<y)$.

Основними кореляційними характеристиками є коефіцієнт кореляції:

$$r = \frac{M\left[\,(X - MX)(Y - MY)\,\right]}{\sigma_X \sigma_Y}, \qquad (3.6.1)$$

лінії регресії Y на x і X на y (лінії умовних математичних сподівань):

$$M(Y \mid X = x), \;\; M(X \mid Y = y). \qquad (3.6.2)$$

Розглянемо вибірку, яка являє собою n пар виду $(x_i, y_j),\ i = 1,...,k,\ j = 1,...,m$.

Рівняння $\overline{Y}_x = f(x)$ називається рівнянням регресії Y на x, де $\overline{Y}_x$ -умовна середня, яка визначається як середнє арифметичне значень Y, які відповідають значенню X=x. Функцію $f(x)$ називають регресією Y на x, а її графік - лінією регресії Y на x.

Аналогічно визначаються умовна середня $\overline{X}_y$ і кореляційна залежність X на y: $\overline{X}_y = \psi(y)$.

Рівняння регресії при заданому вигляді функцій $f(x)$ або $\psi(y)$ знаходять методом найменших квадратів. Так, наприклад, вважаючи функцію $f(x)$ лінійною $f(x) = ax + b$, маємо рівняння прямої лінії регресії Y на x у вигляд:

$$\overline{Y}_x - \overline{y} = r_в \frac{\sigma_y}{\sigma_x}(x - \overline{x}),$$

де $a = r_в \frac{\sigma_y}{\sigma_x},\quad b = \overline{y} - r_в \frac{\sigma_y}{\sigma_x}\overline{x}$, $r_в = \frac{\overline{x} \cdot \overline{y} - \overline{xy}}{\sigma_x \sigma_y}$ - вибірковий коефіцієнт кореляції,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i,\quad \overline{y} = \frac{1}{n}\sum_{j=1}^{m} n_j y_j,\quad \overline{xy} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij} x_i y_j;$$

n_i -частота появи значення x_i; n_j -частота появи значення y_j; n_{ij} - частота пари (x_i, y_j) двовимірної вибірки $(x_i, y_j), i = 1,...,k, j = 1,...,m$; $\overline{x}, \overline{y}, \overline{xy}$ - середні значення компонент x, y та змішаний початковий момент другого порядку,

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - \overline{x})^2,\quad \sigma_y^2 = \frac{1}{n}\sum_{j=1}^{m} n_j (y_j - \overline{y})^2$$

-вибіркові середньоквадратичні відхилення.

Якщо припустити, що функція $\psi(y)$ лінійна, то одержимо рівняння прямої лінії регресії X на y, яка має вигляд:

$$\overline{X}_y - \overline{x} = r_в \frac{\sigma_x}{\sigma_y}(y - \overline{y}).$$

За припущенням, що функція $f(x)$ або $\psi(y)$ зображується кривою лінією, одержуємо рівняння криволінійної кореляції, які також знаходяться методом найменших квадратів.

Приклад 1.

За даними, що наведені у кореляційній таблиці, знайти вибіркове рівняння прямої лінії регресії у на х та оцінити силу кореляційного зв'язку.

x / y	x1	x2	x3
y1	n1	n2	
y2		n3	
y3		n4	n5
y4			n6

Рішення. Зробимо розрахункову таблицю для обчислення коефіцієнта кореляції rB(x,y), який дорівнює:

$$r_B(x,y) = \frac{\sum n_{xy}xy - n\overline{x}\overline{y}}{n\sigma_x\sigma_y},$$

де $n = \sum_{i=1}^{6} n_i$, $\overline{x} = \frac{\sum_{i=1}^{3} x_i}{3}$, $\overline{y} = \frac{\sum_{i=1}^{4} y_i}{4}$, $\overline{x}$ – середнє значення х, $\overline{y}$ – середнє значення у, σх, σу – дисперсія відповідно х та у.

$$\sigma_x = \sqrt{Dx} = \sqrt{\overline{x^2} - \overline{x}^2} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2}{3} - \left(\frac{x_1 + x_2 + x_3}{3}\right)^2},$$

$$\sigma_y = \sqrt{Dy} = \sqrt{\overline{y^2} - \overline{y}^2} = \sqrt{\frac{y_1^2 + y_2^2 + y_3^2 + y_4^2}{4} - \left(\frac{y_1 + y_2 + y_3 + y_4}{4}\right)^2},$$

$\sum n_{xy}xy$ знаходимо за таблицею:

x / y	x1	x2	x3	X	yX
y1	n1x1 n1 n1y1	n2x2 n2 n2y1		n1x1+ n2x2	y1 (n1x1+ n2x2)
y2		n3x2 n3 n3y2		n3x2	y2 n3x2
y3		n4x2 n4	n5x3 n5 n5y3	n4x2+ n5x3	y3 (n4x2+ n5x3)
y4			n6x3 n6 n6y4	n6x3	
Y	n1y1	n2y1+n3y2+n4y3	n5y3+ n6y4		y1 (n1x1+ n2x2)+ +y2 n3x2+ +y3 (n4x2+ n5x3)+ + y4 n6x3
xY	x1 n1y1	x2(n2y1+n3y2+ +n4y3)	x3(n5y3+ n6y4)	x1 n1y1+ + x2(n2y1+n3y2+ n4y3)+ x3(n5y3+ n6y4)	↓ → Σnxyxy

Рівняння лінії регресії має вигляд:

$$\overline{y}_x = r_B(x, y)\frac{\sigma_y}{\sigma_x}(x - \overline{x}) + \overline{y}.$$

Задачі для самостійної роботи.

3.6.1. По заданим статистичним даним розподілу вибірки побудувати гістограму, на підставі якої зробити припущення про вид закону розподілу (нормальний, показниковий, рівномірний). Перевірити припущення за критерієм χ^2. Прийняти рівень значущості $\alpha = 0,05$

Інтер	4-6	6-8	8-10	10-	12-	14-	16-	18-	20-		
Часто	6	9	15	23	30	19	10	7	3		
x_i	0.3	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9	2.1	2.3
2	6	9	26	25	30	26	24	21	20	8	9
3	2-4	4-6	6-8	8-10	10-	12-	14-	16-	18-		
	3	9	25	36	22	17	9	5	3		
4	0-5	5-10	10-	15-	20-	25-	30-	35-	40-		
	3	7	14	26	40	20	13	5	2		
5	(-	(-	(0,1	(10,	(20,	(30,	(40,	(50,			
	21	49	81	90	40	21	10	5			
6	(-	(-	(0,5)	(5,1	(10,	(15,	(20,	(25,	(30,		
	7	12	26	40	70	35	21	10	5		
7	(-5,-	(-3,-	(-	(1,3)	(3,5)	(5,7)	(7,9)	(9,1	(11,		
	7	17	35	47	9	40	24	17	6		
8	5-10	10-	15-	20-	25-	30-	35-	40-	45-		
	7	8	15	18	23	19	14	10	6		
9	6-16	16-	26-	36-	46-	56-	66-	76-	86-		
	8	7	16	35	56	36	17	9	4		
10	1-3	3-5	5-7	7-9	9-11	11-	13-	15-	17-	19-	
	2	4	6	10	18	20	16	11	7	6	
α=0.0	(0,5)	(5,1	(10,	(15,	(20,	(25,	(30,				
11	135	47	17	5	3	2	1				
α=0.0	(0,4	400-	800-	120	160	200	240				

12	121	95	76	56	45	36	21				
α=0.0	0-10	10-	20-	30-	40-	50-	60-				
13	365	245	150	100	70	45	25				
14	0-7	7-14	14-	21-	28-	35-	42-	49-			
	390	274	156	95	63	24	12	7			
15	0-4	4-8	8-12	12-	16-	20-	24-	28-			
	500	324	127	56	42	25	12	6			
16	0-6	6-12	12-	18-	24-	30-	36-				
	390	245	127	75	25	14	5				
17	0-5	5-10	10-	15-	20-	25-	30-				
	425	250	120	50	27	17	6				
18	0-3	3-6	6-9	9-12	12-	15-	18-				
	277	112	56	34	25	12	5				
19	0-4	4-8	8-12	12-	16-	20-	24-				
	370	250	102	37	17	10	3				
20	0-7	7-14	14-	21-	28-	35-	42-				
	517	320	205	114	56	13	8				
21	2	4	6	8	10	12	14	16	18		
	20	16	14	25	21	28	19	24	22		
22	1	5	9	13	17	21	25	29	33		
	31	27	29	28	30	26	25	24	28		
23	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8			
	17	20	16	21	19	17	21	19			
24	(-	(3,7)	(7,1	(11,	(15,	(19,	(23,	(27,			
	10	9	12	13	14	9	10	8			
25	(-	(0,5)	(5,1	(10,	(15,	(20,	(25,				
	31	27	29	26	24	32	23				
26	30	35	40	45	50	55	60	65	70	75	
	4	6	8	15	25	20	8	7	5	2	
27	122	128	134	140	146	152	158	164	170	17	
	7	8	12	16	17	20	13	10	7	3	
28	0-	300-	600-	900-	120	150	180	210	240		
	527	304	217	112	52	24	12	5	2		
29	-4	-2	0	2	4	6	8	10			

	42	45	37	42	27	31	29	40			
30	48	52	56	60	64	68	72	76	80	84	
	2	4	6	8	12	30	18	8	7	5	

3.6.2. Обчислити коефіцієнти кореляції і знайти рівняння прямих регресії Y на x та X на y за даними кореляційної таблиці.

1.

y \ x	40-50	50-60	60-70	70-80
10-11	2	11	3	2
11-12	1	19	2	4
12-13	3	6	27	6
13-14	2	3	3	8

2.

y \ x	5-15	15-25	25-35	35-45	45-55	55-65
10-20	5	7	0	0	0	0
20-30	0	20	23	0	0	0
30-40	0	0	30	47	2	0
40-50	0	0	10	11	20	6
50-60	0	0	0	9	7	3

3.

y \ x	7,0-7,2	7,2-7,4	7,4-7,6	7,6-7,8	7,8-8,0
2,15-2,45	5	4	0	0	0
2,45-2,75	0	12	8	1	0
2,75-3,05	0	0	5	5	0
3,05-3,35	0	0	4	7	0
3,35-3,65	0	0	0	12	1
3,65-3,95	0	0	0	0	1

3.6.3. Знайти довірчі інтервали для математичного сподівання і середнього квадратичного відхилення нормального розподілу з надійністю 0,95, знаючи вибіркову середню $\overline{x}$, об'єм вибірки n, вибіркову дисперсію s2.

N	$\overline{x}$	n	s2	N	$\overline{x}$	N	s2
1	75,17	21	7	16	12,5	9	5
2	65,20	22	6	17	21,7	12	7
3	15,13	20	9	18	31,4	17	6
4	25,21	21	10	19	28,5	19	10
5	17,40	18	11	20	37,11	21	11
6	27,14	17	12	21	42,12	23	12
7	45,12	19	11	22	11,7	8	4
8	48,17	18	13	23	15,7	9	5
9	50,12	22	15	24	21,7	12	6
10	65,07	17	12	25	31,12	15	8
11	57,08	22	14	26	5	16	4
12	59,00	21	10	27	61,3	21	13
13	25,14	17	9	28	28,4	17	11
14	37,12	19	17	29	31,5	14	8
15	49,14	17	10	30	39,12	15	7

Литература

1. Венцель Е.С. Теория вероятностей.–М.: Наука, 1969.
2. Гмурман В.Е. Руководство к решению задач по теории вероятности и математической статистике.–М.: «Высшая школа», 1975.
3. Гмурман В.Е. Теория вероятности и математическая статистика. –М.: «Высшая школа», 1977.
4. Чистяков В.П. Курс теории вероятностей.– М.: Наука, 1982.
5. Колемаев В.А. и др. Теория вероятностей и математическая статистика. – М: Высш. шк., 1991.

Таблиця значень функції $\varphi = \frac{1}{\sqrt{2\pi}} \ell^{\frac{x^2}{2}}$

	0	1	2	3	4	5	6	7	8	9
0,0	0, 3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3647
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	§332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0, 2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551

РЕЦЕНЗІЯ

на навчальний посібник „Теорія ймовірностей і математична статистика” авторів Говаленкова С.В., Комяк В.М., Мігунової Л.В., Тарасенко О.А.

В результаті розгляду навчального посібника „Теорія ймовірностей і математична статистика” авторів Говаленкова С.В., Комяк В.М., Мігунової Л.В., Тарасенко О.А. слід зазначити наступне:

1. Передбачений для вивчення учбовий матеріал поділено на 3 розділи: випадкові події, випадкові величини, математична статистика. Прийнята послідовність викладення матеріалу логічно узгоджена та достатньо обґрунтована, узгоджується з тим, що потрібно слухачам для успішного використання математичних методів в подальшому вивченні спеціальних дисциплін і при вирішенні прикладних задач пожежної охорони.
2. До параграфів розділів наведені приклади вирішення задач, що сприятиме більш якісному освоєнню учбового матеріалу.
3. Викладення матеріалу узгоджується з прийнятою НМК „Пожежна безпека” навчальною програмою з дисципліни „Вища математика”.
4. З метою сприяння якісній самостійній роботі над предметом рекомендується досить повний перелік літератури, яка відображає сучасний стан навчальної літератури з дисципліни і охоплює сучасні та минулі видання.

Вважаю, що запропонований навчальний посібник „Теорія ймовірностей і математична статистика” може бути рекомендований до друку та запровадження в учбовий процес вищих навчальних закладів освіти України пожежно-технічного профілю.

Начальник кафедри «Прикладної механіки»
Д.ф.-м.н., проф. В.П.Ольшанський

ЗМІСТ