

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ
МІСЬКОГО ГОСПОДАРСТВА**

**А. Є. Ачкасов, О. О. Воронков,
Т. Б. Воронкова**

КОНСПЕКТ ЛЕКЦІЙ

з курсу

«ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА»

(для студентів 2 курсу ФПО і ЗН напряму підготовки
(0921) 6.060101 «Будівництво» та слухачів другої вищої освіти
за спеціальністю «Теплогазопостачання і вентиляція»)

ХАРКІВ ХНАМГ 2007

Ачкасов А. Є. Конспект лекцій з курсу «Теорія ймовірностей і математична статистика» (для студентів 2 курсу ФПО і ЗН напряму підготовки (0921) 6.060101 «Будівництво» та слухачів другої вищої освіти за спеціальністю «Теплогазопостачання і вентиляція») / А. Є. Ачкасов, О. О. Воронков, Т. Б. Воронкова; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2007. – 103 с.

Автори: А. Є. Ачкасов,
О. О. Воронков,
Т. Б. Воронкова

Рецензент: доц. С. О. Станішевський

Рекомендовано кафедрою «Економіки й управління в будівництві і міському господарстві»,
протокол № 1 від 28.08.07 р.

ВСТУП

Курс «Теорія ймовірностей і математична статистика» є нормативною дисципліною в навчальному плані за напрямком «Економіка й підприємництво» для кваліфікаційного рівня «Бакалавр». Обсяг курсу становить 108 академічних годин або три кредити. При вивченні за заочною формою обсяг аудиторних занять становить 14 годин (8 годин лекцій та 6 годин практичних занять), на самостійну роботу студента залишається 94 години. Програма курсу включає три змістових модуля: «Теорія ймовірностей», «Математична статистика» і «Випадкові процеси», відповідно до яких виконується проміжний контроль знань. Підсумковий контроль знань (залік) проводиться в усній формі. У процесі вивчення курсу студенти повинні виконати контрольну роботу.

Метою вивчення дисципліни «Теорія ймовірностей і математична статистика» є формування базових знань в області застосування ймовірностатистичного апарата, вивчення закономірностей у масових випадкових явищах, визначення їхніх імовірнісних характеристик з метою прогнозування.

У результаті вивчення курсу студенти повинні оволодіти основними методами визначення імовірнісних характеристик випадкових величин, статистичного опису результатів спостереження й перевірки статистичних гіпотез із метою прийняття на їхній основі обґрунтованих рішень.

Теорія ймовірностей і математична статистика – математична наука, що вивчає закономірності у випадкових явищах. Випадковим називається таке явище, яке при багаторазовому повторенні досліду протікає щораз по-іншому. Виникає питання, чи у кожному випадковому явищі міститься закономірність? Якщо випадкове явище має так звану статистичну однорідність, то воно містить закономірність і називається стохастичним. У процесі своєї діяльності людина часто стикається з випадковістю й інтуїтивно припускає наявність у цій випадковості закономірності. Однак іноді покладатися на інтуїцію небажано, все залежить від складності й важливості розв’язуваної проблеми. Тоді виникає необхідність визначити ступінь можливості яких-небудь подій або наслідків шля-

хом математичних розрахунків. Тут і доводиться звертатися до методів теорії ймовірностей і математичної статистики.

Зауважимо, що економічні процеси й події є у певній мірі випадковими. Причому, це пов'язано не тільки з «чистою» випадковістю й довільним поведінням людини як елемента економічної системи, а й зі складністю процесів, які піддані впливу такої множини факторів, що не завжди можна врахувати їхній вплив цілком. У зв'язку із цим застосування ймовірнісно-статистичного апарата в економічних розрахунках виявляється дуже актуальним і ефективним. Зокрема, при плануванні й прогнозуванні або при оцінці ризику при вкладенні інвестицій або реалізації бізнес-плану.

Інформація про випадкове явище може бути отримана в результаті його спостереження, тобто шляхом проведення дослідів. Для виявлення закономірності провадять обробку дослідних даних. Цю задачу вирішує математична статистика. Теоретичною базою математичної статистики є класична теорія ймовірностей.

Початок розвитку теорії ймовірностей і математичної статистики пов'язаний з європейськими математиками XVII століття. Завдяки роботам швейцарського математика Якоба Бернуллі теорія ймовірностей набула найважливішого значення в практичній діяльності. Він побудував математичну модель для опису серії незалежних випробувань, довів теорему, що є особливим випадком закону великих чисел (теорему Бернуллі), що має основне значення в теорії ймовірностей і її застосування до математичної статистики. В XVIII столітті англійський математик Томас Бейес поставив і вирішив одну з основних задач елементарної теорії ймовірностей – теорему гіпотез, відому за назвою «формула Бейеса». Французький математик П'єр Симон Лаплас розвив і систематизував результати, отримані Бернуллі. Він довів важливу граничну теорему (теорему Лапласа-Муавра), розвив теорію помилок, обґрунтував, хоча й нестрого, метод найменших квадратів. Теорія ймовірностей у значній мірі сформувалася саме в його роботах. Він запровадив теореми додавання й множення ймовірностей, поняття математичного сподівання й виробляючих функцій. За його життя

робота «Аналітична теорія ймовірностей» видавалася тричі. Учень Лапласа, французький математик Сімеон Дені Пуассон ґрунтовно розвив ідеї Лапласа. Він довів теорему, що стосувалася закону великих чисел (закон Пуассона), вперше скориставшись терміном «закон великих чисел».

В XIX столітті теорія ймовірностей сформувалася як злагоджена математична дисципліна в зв'язку з видатними роботами російського математика Пафнутія Львовича Чебишева і його учнів Ляпунова О.М. і Маркова А.А. Чебишев П.Л. довів загальні форми закону великих чисел. Марков Андрій Андрійович збагатив теорію ймовірностей важливими відкриттями й методами. Він розвив метод моментів Чебишева настільки, що став можливим доказ центральної граничної теореми; істотно розширив сферу застосування закону великих чисел і центральної граничної теореми, поширивши їх не тільки на незалежні, але й на залежні дослідження; заклав основи однієї із загальних схем природних процесів, що згодом назвали ланцюгами Маркова. Це привело до розвитку нового розділу теорії ймовірностей – теорії випадкових процесів. У математичній статистиці А.А.Марков вивів принцип, еквівалентний поняттям незміщених і ефективних статистик. Ляпунов Олександр Михайлович працював у Харківському університеті. Він зробив важливий внесок у теорію ймовірностей, давши простий і строгий доказ центральної граничної теореми в більш загальній формі, ніж та, в якій вона розглядалася Чебишевим і Марковим. Для доказу він розробив метод характеристичних функцій, що широко застосовується в сучасній теорії ймовірностей.

У XX столітті значний внесок у розвиток сучасної теорії ймовірностей внесли радянські вчені. Колмогоров Андрій Миколайович разом з А.Я.Хінциним вирішив широко відому систему аксіоматичного обґрунтування, створив теорію марковських процесів з безперервним часом, розвив теорію стаціонарних випадкових процесів. Колмогорову А.М. належать дослідження зі статистичних методів контролю масової продукції й теорії передачі інформації каналами зв'язку. Смирнов Микола Васильович отримав фундаментальні результати з розподілу елементів варіаційного ряду й інших питань теорії ймовір-

ностей і математичної статистики. У теорії граничних теорем відомий критерій Смирнова. Гнеденко Борис Володимирович вирішив питання щодо умов збіжності розподілу сум незалежних доданків до всіх можливих для них розподілів, отримав важливі результати з теорії масового обслуговування й теорії надійності.

Застосування імовірнісних і статистичних методів дає можливість вивчати на науковій основі як діяльність окремих підприємств, так і соціально-економічні процеси в суспільстві в цілому.

Теорія ймовірностей і математична статистика є основою для побудови кількісних моделей керування економічними системами. Прикладом таких моделей служать моделі планування й керування запасами, теорії ігор, теорії масового обслуговування. Ймовірнісно-статистичні методи є базовими для теорії прийняття рішень – складової частини сучасного менеджменту. Статистичні показники аналізують при оцінці ризику в інвестиційній діяльності, в діяльності страхових компаній, а також у багатьох сферах економіки й керування.

Широка сфера застосування теорії ймовірностей і математичної статистики визначає важливе місце, що займає даний курс у підготовці економістів вищої кваліфікації, а також менеджерів, маркетологів, бізнесменів, бухгалтерів-аудиторів, соціологів. Ця дисципліна закладає теоретичну основу для наступного вивчення курсів «Математичне програмування», «Економетрія», «Статистика», «Економічний аналіз», «Економічні ризики», «Теорія прийняття рішень».

Змістовий модуль 1

ІМОВІРНІСТЬ ВИПАДКОВОЇ ПОДІЇ

Тема 1. Основні поняття теорії ймовірностей

Теорія ймовірностей – математична наука, що вивчає закономірності у випадкових явищах. Випадковим називається таке явище, яке при багаторазовому повторенні досліду протікає щораз по-іншому. Наприклад, вимірювання – якщо ми бажаємо отримати точний результат, стрілянина по мішені являє класичний приклад випадкового явища, погодні умови і т. інше.

На відміну від випадкових явищ існують детерміновані явища. Це, як правило, закони природи, які вивчаються в курсі фізики, наприклад:

- прискорення вільного падіння дорівнює $9,8 \text{ м/с}^2$;
- сила, прикладена до матеріальної точки, надає їй прискорення (за формулою $F=ma$);
- струм, що протікає через опір R під дією напруги U , пропорційний напрузі $I=U/R$.

Таким чином, якщо при відтворенні певних умов незмінно відбувається певна подія (та сама, тобто результат незмінно повторюється), то має місце детерміноване явище. Прогноз результату такого досліду можна здійснити, не проводячи експерименту.

У випадку, коли на результат досліду впливає ряд факторів, урахувати які неможливо або дуже складно (крім того, ці фактори можуть змінюватися випадковим образом), скласти математичну модель, що прогнозує розвиток такого явища в детерміністському поданні неможливо. У такому випадку намагаються знайти у випадкових явищах ті або інші закономірності. Такі закономірності виявляються при масовому повторенні дослідів. Якщо їх удається знайти, то випадкове явище є статистично однорідним. Якщо закономірності в явищі від-

сутні, тобто виявити їх не вдається, то таке явище є невизначеним і вимагає додаткового дослідження.

Одним з основних у теорії ймовірностей є поняття випадкової події.

Випадкова подія – це всякий факт, що в результаті досліду може відбутися або не відбутися.

Випадкові події позначають великими буквами латинського алфавіту: $A = \{\text{влучення в мішень}\}$, $B = \{\text{прибуття трамвая на зупинку}\}$, $C = \{\text{поломка технічного устрою}\}$, $D = \{\text{коротке замикання в мережі}\}$.

Дослідом називається відтворена сукупність умов, у яких може відбутися випадкова подія.

Імовірність випадкової події – це числова міра ступеню об'єктивної можливості появи даної події в результаті досліду. Імовірність події A позначається $P(A)$.

За одиницю виміру ймовірності приймають імовірність достовірної події E , тобто такої, яка в результаті досліду обов'язково відбудеться:

$$P(E) = 1.$$

Протилежна достовірному подія називається неможливою і позначається \bar{E} . Імовірність неможливої події:

$$P(\bar{E}) = 0.$$

Відповідно ймовірність будь-якої випадкової події A укладена між нулем і одиницею:

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

Класичний і статистичний методи визначення ймовірності випадкової події

Імовірність випадкової події можна визначити класичним методом тільки в обмеженому числі явищ, а саме, якщо наслідки досліду мають наступні властивості:

утворюють повну групу – якщо результатом однократного випробування є обов'язково один з можливих наслідків;

є рівноможливими – якщо за умовами симетрії досліду поява кожного з них однаково можлива;

є несумісними – якщо будь-які два з них не можуть відбутися одночасно.

Якщо наслідки досліду мають перелічені властивості (утворюють повну групу, є несумісними й рівноможливими), то говорять, що дослід зводиться до **схеми випадків**, або що має місце класична схема теорії ймовірностей. У рамках цієї схеми можна точно підрахувати імовірність події, не проводячи випробувань. Якщо дослід зводиться до схеми випадків, то імовірність події визначається як відношення числа можливих наслідків досліду, що сприяють появі події А, до загального числа можливих наслідків досліду

$$P(A) = \frac{m}{n}, \quad (1.1)$$

де n – загальне число можливих наслідків досліду; m – число наслідків досліду, що сприяють появі події А.

Для підрахунку числа всіх випадків n і числа випадків m , що сприяють появі події А, часто використовують число сполучень із k елементів по s

$$C_k^s = \frac{k!}{s!(k-s)!}$$

де $k! = 1*2*3 \dots *k$, при цьому $0! = 1$.

Якщо події в досліді не зводяться до схеми випадків, то оцінку імовірності можна зробити тільки статистично. Спостерігаючи випадкові явища або проводячи випробування, визначають частоту появи даної події. При проведенні серії з n дослідів, у кожному з яких могла з'явитися або не з'явитися подія А, під частотою її появи розуміють відношення

$$P^*(A) = \frac{m}{n} \quad (1.2)$$

де n – число проведених дослідів; m – число появ події А в n дослідях.

Чи можна вважати частоту появи події A її імовірністю? Результат кожного досліду є випадковим, однак якщо спостережуване явище має статистичну однорідність, то при великій кількості дослідів частота події починає стабілізуватися й у межі прагне до ймовірності події. Це властивість стійкості частот, багаторазово перевірена експериментально, і є одна з найбільш характерних закономірностей, спостережуваних у випадкових явищах. Вона відома за назвою закону великих чисел. Бернуллі довів, що

$$\lim_{n \rightarrow \infty} P^*(A) = P(A). \quad (1.3)$$

Вираз (1.3) читається так: імовірність події A із збільшенням числа дослідів n сходиться за імовірністю до імовірності події A . Це означає, що зі збільшенням числа дослідів n імовірність, що частота події A відрізняється від імовірності цієї події зменшується.

Тема 2. Операції над подіями. Теорема теорії ймовірностей. Основні формули теорії ймовірностей

Оскільки в практичних умовах багаторазове відтворення досліду надзвичайно утруднено, для визначення ймовірностей одних випадкових подій по відомих ймовірностях інших подій, з ними пов'язаних, користуються теоремами теорії ймовірностей: теоремою додавання й теоремою добутку.

Введемо визначення.

Сумою двох подій A і B називається подія C , що перебуває в появі події A або події B або обох разом:

$$C = A + B.$$

Сума подій – логічна сума, вона називається диз'юнкцією й позначається спеціальним знаком:

$$C = A \cup B.$$

Добутком двох подій A і B називається подія C , що перебуває в спільній появі подій A і B :

$$C = A * B.$$

Добуток подій – логічний добуток, називається кон'юнкцією і також позначається спеціальним знаком:

$$C = A \cap B.$$

Протилежними називаються дві несумісних події A і \bar{A} , якщо вони складають повну групу.

Подія A називається **незалежною** від події B , якщо імовірність події A не змінюється від того, відбулася подія B чи ні. Якщо ж імовірність події A залежить від того, відбулася подія B чи ні, то такі події називаються **залежними**.

Імовірність події A , обчислена за умови, що подія B мала місце, називається **умовною ймовірністю** події A і позначається $P(A|B)$.

Розглянемо приклад. Нехай в урні три кулі, дві з яких білі, а третя – чорна. Одну за іншою з урни виймають дві кулі. Позначимо події

$$A = \{\text{перша вийнята куля виявилася білою}\}$$

$$B = \{\text{друга вийнята куля виявилася білою}\}.$$

Імовірність події B залежить від того, відбулася подія A чи ні. Якщо подія A відбулася, то імовірність події B

$$P(B/A) = \frac{1}{2}.$$

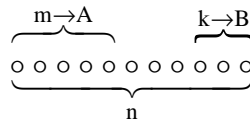
Якщо ж подія A не відбулася, то імовірність події B буде іншою: якщо першою виявилася вийнятою чорна куля, то $P(B/\bar{A}) = 1$.

Теорема додавання

Імовірність суми двох несумісних подій A і B дорівнює сумі ймовірностей цих подій, тобто

$$P(A + B) = P(A) + P(B) \quad (2.1)$$

Доведемо це. Нехай дослід має n можливих наслідків, в числі яких m сприяють появі події A і k – появі події B .



Оскільки подія С перебуває в появі події А, якої сприяє m наслідків досліду або події В, якої сприяє k наслідків, то події С сприяють $m+k$ наслідків досліду. Тоді імовірність події С за класичною формулою визначиться в такий спосіб:

$$P(C) = \frac{m+k}{n} = \frac{m}{n} + \frac{k}{n} = P(A) + P(B).$$

Слідства теореми додавання:

За методом математичної індукції (узагальнення) теорему додавання ймовірностей можна розповсюдити на будь-яке кінцеве число несумісних подій:

$$C = \Sigma A_i$$

$$P(\Sigma A_i) = \Sigma P(A_i).$$

Слідство 1. Якщо події $A_1, A_2, \dots, A_i, \dots, A_n$ утворюють повну групу несумісних подій, то сума їхніх імовірностей дорівнює 1:

$$P(\Sigma A_i) = \Sigma P(A_i) = 1. \quad (2.2)$$

Слідство 2. Сума ймовірностей двох протилежних подій дорівнює одиниці:

$$P(A) + P(\bar{A}) = 1,$$

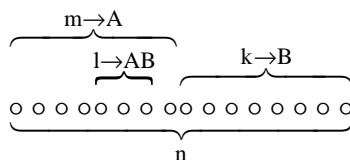
звідки імовірність будь-якої випадкової події:

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

У випадку, коли дві події є сумісними, імовірність їхньої суми дорівнює сумі ймовірностей цих подій мінус імовірність їхньої спільної появи:

$$P(A + B) = P(A) + P(B) - P(A*B) \quad (2.3)$$

Доведемо це. Нехай дослід має n можливих наслідків, в числі яких m сприяють появі події А, k – появі події В і l – появі події АВ.



Оскільки подія С перебуває в появі події А, якої сприяє m наслідків досліду або події В, якої сприяє k наслідків, то події С сприяють $m+k-1$ наслідків досліду. Тоді імовірність події С за класичною формулою визначиться в такий спосіб:

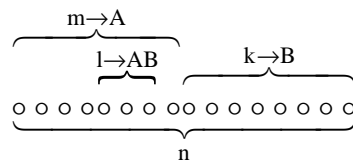
$$P(C) = \frac{m+k-1}{n} = \frac{m}{n} + \frac{k}{n} - \frac{1}{n} = P(A) + P(B) - P(AB).$$

Теорема множення

Імовірність добутку двох подій А і В дорівнює добутку ймовірності одного з них на умовну ймовірність іншого, обчислену за умови, що перша відбулася:

$$P(A * B) = P(A) * P(B|A). \tag{2.4}$$

Доведемо це. Нехай дослід має n можливих наслідків, в числі яких m сприяють появі події А, k – появі події В і l – появі події АВ.



Події АВ, сприяють l наслідків досліду:

$$P(AB) = \frac{l}{n}; \quad P(A) = \frac{m}{n}; \quad P(B/A) = \frac{l}{m}; \quad P(B) = \frac{k}{n}; \quad P(A/B) = \frac{l}{k}.$$

Тоді імовірність події АВ визначиться в такий спосіб:

$$P(AB) = \frac{l}{n} = \frac{m}{n} * \frac{l}{m} = P(A) * P(B/A).$$

Якщо події А і В незалежні, то умовна ймовірність події В дорівнює безумовній імовірності цієї події,

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

Слідство. Імовірність добутку двох незалежних подій дорівнює добутку ймовірностей цих подій:

$$P(A * B) = P(A) * P(B) \tag{2.5}$$

Якщо маємо кілька незалежних подій:

$$P\left(\prod_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i).$$

Формула повної ймовірності

Формула повної ймовірності є слідством двох теорем теорії ймовірностей. Нехай передбачається проведення досліду, про умови протікання якого можна зробити N взаємовиключних припущень (гіпотез). Умови протікання досліду (гіпотези) являють собою повну групу неспільних подій H_1, H_2, \dots, H_N , імовірності яких $P(H_i)$ відомі. Деяка випадкова подія A може з'явитися при будь-яких умовах протікання досліду з різною ймовірністю. Уявимо подію A як суму не-сумісних подій:

$$A = H_1A + H_2A + \dots + H_NA.$$

Застосовуючи теорему додавання й множення, дістанемо:

$$P(A) = \sum_{i=1}^N P(H_iA) = \sum_{i=1}^N P(H_i)P(A/H_i). \quad (2.6)$$

Таким чином, повна безумовна ймовірність події A з урахуванням випадковості умов протікання досліду дорівнює сумі добутків ймовірностей кожної з гіпотез на умовну ймовірність події A при кожній з гіпотез.

Формула Бейеса (теорема гіпотез)

Ця формула дозволяє за відомими до проведення досліду (ап'юріорним) ймовірностями гіпотез $P(H_i)$ і за результатом досліду (наступ події A) визначити обчислені після досліду (апостеріорні) ймовірності гіпотез $P(H_i|A)$.

За теоремою множення ймовірність появи події A при i -й гіпотезі

$$P(H_i * A) = P(H_i) * P(A|H_i)$$

у чинність симетрії подій справедливо

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

звідки дістанемо

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

або, якщо підставити $P(A)$ з формули (2.6), дістанемо

$$P(H_i / A) = \frac{P(H_i) * P(A / H_i)}{\sum P(H_i) * P(A / H_i)} \quad (2.7)$$

Таким чином, формула Бейеса дозволяє переоцінити імовірності гіпотез після того, як стає відомим результат дослід, в результаті якого відбулася подія A .

Формула Бернуллі

На практиці доводиться зіштовхуватися з такими задачами, які можна представити у вигляді багаторазово повторюваних незалежних випробувань. Якщо імовірність появи події A в одному досліді та сама, то імовірність m появ події A в n дослідях можна визначити за формулою Бернуллі

$$P_n(m) = C_n^m p^m q^{n-m} \quad (2.8)$$

де $P_n(m)$ - імовірність того, що в n випробуваннях подія A з'явиться рівно m разів;

C_n^m - число сполучень із n елементів по m ;

p - імовірність появи події A в одному досліді;

$q = 1 - p$ - імовірність не появи події A в одному досліді.

Користуючись формулою Бернуллі, можна дістати найімовірніше число появ події A .

Найімовірнішим числом появи події A в n незалежних дослідях називається таке число, для якого імовірність перевищує або, принаймні, не менше ймовірності кожного з інших можливих чисел появи події A . Для визначення найімовірнішого числа користуються формулою:

$$np - q \leq m_0 \leq np + p \quad (2.9)$$

причому, m_0 — може бути тільки цілим числом. Якщо $np - q$ - ціле число, то $m_0 = np$.

Локальна теорема Лапласа

Локальна теорема Лапласа дає асимптотичну формулу, що дозволяє при-

близно знайти імовірність появи подій рівно m разів в n дослідах, якщо число випробувань досить велике.

Якщо імовірність p появи події A в кожному випробуванні постійна й відмінна від нуля й одиниці, то імовірність $P_n(m)$ того, що подія A з'явиться в n дослідах рівно m разів, приблизно дорівнює, і тим точніше, чим більше n , значенню функції

$$P_n(m) = \frac{1}{\sqrt{npq}} \varphi(x), \quad (2.10)$$

де $x = \frac{m - np}{\sqrt{npq}}$, а значення функції $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ визначаються з довідкових

таблиць. Функція $\varphi(x)$ парна, тобто $\varphi(-x) = \varphi(x)$.

Інтегральна теорема Лапласа

Якщо в повторних незалежних випробуваннях, у кожному з яких імовірність появи події A постійна й дорівнює p , необхідно обчислити ймовірність того, що подія A з'явиться в n випробуваннях не менше m_1 і не більше m_2 разів, це можна зробити за допомогою інтегральної теореми Лапласа.

Якщо імовірність p настання події A в кожному випробуванні постійна й відмінна від нуля й одиниці, то приблизно імовірність $P_n(m_1, m_2)$, того, що подія A з'явиться у випробуваннях від m_1 до m_2 разів,

$$P_n(m_1, m_2) = \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-x^2/2} dx, \quad (2.11)$$

де $x' = \frac{m_1 - np}{\sqrt{npq}}$; $x'' = \frac{m_2 - np}{\sqrt{npq}}$

При розв'язанні задач користуються спеціальними таблицями, тому що невизначений інтеграл (2.11) не виражається через елементарні функції. У довідкових

таблицях приводяться значення інтеграла $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-x^2/2} dx$. Функція $\Phi(x)$ непарна, тобто $\Phi(-x) = -\Phi(x)$. Таблиця містить значення функції $\Phi(x)$

тільки для $x \in [0; 5]$; для $x > 5$ приймають $\Phi(x) = 0,5$.

Таким чином, приблизно імовірність того, що подія А з'явиться в n незалежних дослідах від m_1 до m_2 разів.

$$P_n(m_1, m_2) = \Phi(x_2) - \Phi(x_1), \quad (2.12)$$

Формула Пуассона

Якщо імовірність p настання події в окремому випробуванні близька до нуля, то навіть при великій кількості випробувань n , але при невеликому значенні добутку np одержувані за формулою Лапласа значення ймовірностей $P_n(m)$ виявляються недостатньо точними й виникає потреба в іншій наближеній формулі.

Якщо імовірність p настання події А в кожному випробуванні постійна, але мала, число незалежних випробувань n досить велике, але значення добутку np залишається невеликим (не більше десяти), то імовірність того, що в цих випробуваннях подія А наступить m разів, можна визначити за формулою Пуассона:

$$P_n(m) = \frac{(np)^m}{m!} e^{-np} \quad (2.13)$$

Змістовий модуль 2

ВИПАДКОВІ ВЕЛИЧИНИ І ЇХНІ ЗАКОНИ РОЗПОДІЛУ

Тема 3. Поняття випадкової величини. Універсальні закони розподілу

Випадковою називається така величина, що в результаті дослідження може прийняти те або інше значення, невідомо заздалегідь, яке саме (наприклад, число очків, число пасажирів у трамваї, температура повітря, час напрацювання на відмову технічного устрою). Випадкову величину позначають прописною літерою латинського алфавіту X , Y або Z , а будь-яке її значення – відповідною малою літерою x , y або z .

Розрізняють дискретні й безперервні випадкові величини.

Дискретною називається така випадкова величина, число значень якої скінченне, або нескінченне, але рахункове (яка може приймати тільки окремі значення).

Безперервною називається така випадкова величина, число значень якої нескінченне навіть на невеликому інтервалі.

Для повної характеристики випадкової величини необхідно знати всі можливі її значення, а також імовірності появи цих значень у результаті дослідження.

Законом розподілу випадкової величини називається будь-яке правило, що дозволяє знаходити ймовірності різних подій, пов'язаних з випадковою величиною, зокрема, імовірність того, що вона прийме якесь значення або потрапить у якийсь інтервал своїх значень.

Найпростішим законом розподілу є закон розподілу дискретної випадкової величини X , називаний *рядом розподілу*. Він являє собою таблицю, у верхньому рядку якої перелічені всі значення випадкової величини x_1, x_2, \dots, x_n у порядку їхнього зростання, а в нижньої – імовірності появи цих значень p_1, p_2, \dots, p_n :

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

де $p_i = P\{X=x_i\}$.

Оскільки події $\{X=x_1\}, \{X=x_2\}, \dots, \{X=x_n\}$ неспільні й утворюють повну групу, сума їхніх ймовірностей дорівнює одиниці $\sum p_i = 1$ (ця одиниця розподілена між значеннями X). Ряд розподілу може бути заданий і у вигляді графіка:

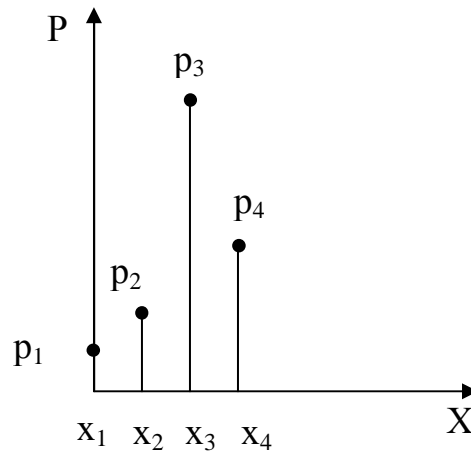


Рис. 3.1 – Графічна інтерпретація ряду розподілу.

Найбільш загальною формою закону розподілу для всіх випадкових величин є функція розподілу.

Функція розподілу

Для характеристики як безперервних так і дискретних випадкових величин зручніше користуватися не ймовірністю події $X=x_i$ (тому що значень x_i може бути багато), а ймовірністю події $X < x$.

Функція розподілу випадкової величини X дорівнює імовірності того, що випадкова величина X прийме значення, менше x :

$$F(x) = P\{X < x\} \quad (3.1)$$

Геометрично функція розподілу – це імовірність того, що значення випадкової величини потрапить лівіше x_i .

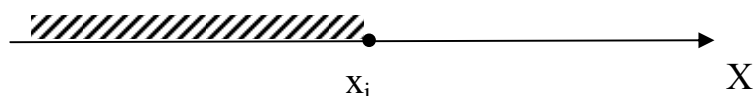


Рис. 3.2 – Розташування значень $X < x$

Функція розподілу дискретної випадкової величини являє собою розривну, східчасту функцію, що має перегони в точках, що відповідають можливим значенням x_1, x_2, \dots, x_n випадкової величини X , які дорівнюють ймовірностям p_1, p_2, \dots, p_n цих значень.

У випадку безперервної випадкової величини функція розподілу звичайно має вигляд плавної кривої.

Щільність розподілу

Нехай ϵ безперервна випадкова величина X з функцією розподілу $F(x)$.

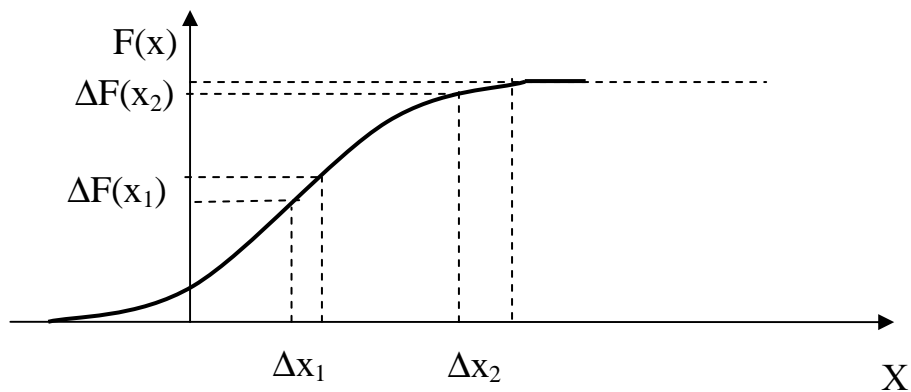


Рис. 3.3 – Функція розподілу безперервної випадкової величини

Говорити про розподіл ймовірностей між значеннями безперервної випадкової величини нема рації, тому що число її значень нескінченно навіть на невеликому інтервалі, й імовірність того, що безперервна випадкова величина прийме одне-єдине своє значення x_i дорівнює нулю. Тому, характеризуючи безперервну випадкову величину, завжди говорять про влучення її значень у той чи інший інтервал. З рисунка 3.3 видно, що імовірність влучення X в інтервал Δx_1 більше, ніж в інтервал Δx_2 , оскільки приріст функції розподілу $\Delta F(x_1) > \Delta F(x_2)$.

З математики відомо, що
$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF(x)}{dx}$$

Таким чином, закон розподілу ймовірностей безперервних випадкових величин зручніше визначати завданням не функції розподілу $F(x)$, а щільності розподілу ймовірностей $f(x)$, що являє собою похідну від $F(x)$ по x :

$$f(x) = \frac{dF(x)}{dx} \quad (3.2)$$

Передбачається, що $F(x)$ безперервна й диференційована.

Щільністю розподілу випадкової величини X у точці x називається похідна функції розподілу X у цій точці.

На графіку щільності розподілу ймовірність представляється площею.

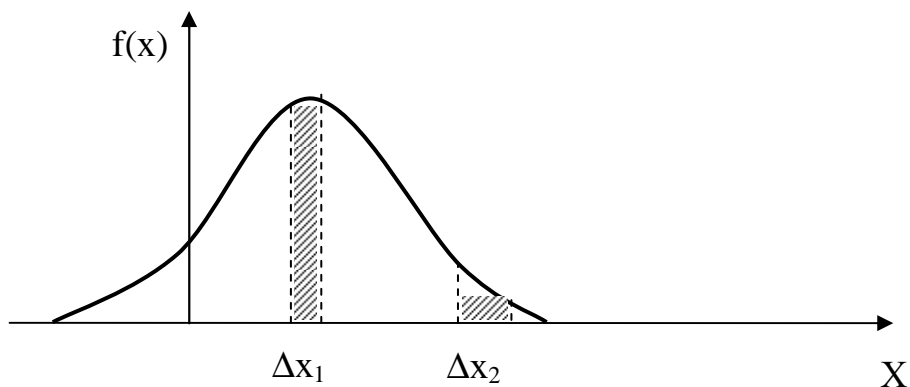


Рис. 3.4 – Графік щільності розподілу

Властивості щільності розподілу ймовірностей:

1. Щільність розподілу невід'ємна, тобто $f(x) \geq 0$ як похідна неубутної функції.
2. Інтеграл від щільності розподілу в нескінченних межах дорівнює одиниці

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

3. Імовірність влучення безперервної випадкової величини в інтервал (x_1, x_2)

$$P\{x_1 \leq X \leq x_2\} = \int_{x_1}^{x_2} f(x)dx .$$

4. Функція розподілу визначається співвідношенням:

$$F(x) = \int_{-\infty}^x f(x) dx .$$

Функцію розподілу $F(x)$ іноді називають інтегральним законом розподілу, а щільність розподілу $f(x)$ – диференціальним законом розподілу.

Тема 4. Числові характеристики випадкової величини

Закон розподілу випадкової величини являє собою деяку функцію, що цілком описує випадкову величину з імовірнісної точки зору, тобто є її вичерпною характеристикою й дозволяє визначати імовірності будь-яких подій, пов'язаних з випадковою величиною. Однак у багатьох практичних задачах потрібно отримати більш компактне уявлення про випадкову величину. Для теорії ймовірностей і її застосування велику роль грають деякі постійні числа, що одержані за певними правилами із законів розподілу випадкових величин і названі **числовими характеристиками** випадкової величини. Найважливішою числовою характеристикою випадкової величини є **математичне сподівання**.

Математичним сподіванням випадкової величини X називається сума добутків всіх можливих її значень на імовірності цих значень

$$M[X] = \sum_{i=1}^n x_i p_i . \quad (4.1)$$

Математичне сподівання характеризує середнє значення випадкової величини, навколо якого групуються всі її можливі значення.

На практиці використовується ряд числових характеристик, які характеризують особливості розподілу випадкової величини. Такими характеристиками є так звані **моменти** або математичні сподівання випадкової величини. Розрізняють початкові α і центральні μ моменти.

Початковим моментом s -го порядку дискретної випадкової величини називається математичне сподівання s -го ступеня цієї величини

$$\alpha_s = \sum_{i=1}^n x_i^s p_i . \quad (4.2)$$

Для безперервної випадкової величини початковий момент s –го порядку

$$\alpha_s = \int_{-\infty}^{+\infty} x^s f(x) dx . \quad (4.3)$$

Вираження (4.2) і (4.3) можна об'єднати в одне, користуючись знаком математичного сподівання M :

$$\alpha_s = M[X^s] , \quad (4.4)$$

тобто початковим моментом s –го порядку випадкової величини називається математичне сподівання s –го ступеня цієї випадкової величини.

При $s = 1$ дістаємо перший початковий момент або математичне сподівання випадкової величини

$$\alpha_1 = M[X] = m_x . \quad (4.5)$$

На практиці іноді застосовують другий початковий момент α_2 :

$$\alpha_2 = M[X^2] \quad (4.6)$$

Центральним моментом s -го порядку випадкової величини X називається математичне сподівання s –го ступеня центрованої величини X . Під центрованою розуміється відхилення випадкової величини від її математичного сподівання:

$$X^0 = X - m_x$$

Центральний момент s -го порядку випадкової величини X виражається формулою

$$\mu_s = M[X^0{}^s] \quad (4.7)$$

Для дискретної випадкової величини X :

$$\mu_s = M[X^s] = \sum_{i=1}^n x_i^s p_i = \sum_{i=1}^n (x_i - m_x)^s p_i \quad (4.8)$$

Для безперервної випадкової величини X:

$$\mu_s = M[X^s] = \int_{-\infty}^{+\infty} x^s f(x) dx = \int_{-\infty}^{+\infty} (x - m_x)^s f(x) dx \quad (4.9)$$

Центральний момент першого порядку дорівнює нулю:

$$M[X] = M[X - m_x] = \sum_{i=1}^n (x_i - m_x) p_i = \sum_{i=1}^n x_i p_i - m_x \sum_{i=1}^n p_i = 0.$$

Центральний момент другого порядку для дискретної випадкової величини X:

$$m_2 = M[X^2] = \sum_{i=1}^n x_i^2 p_i = \sum_{i=1}^n (x_i - m_x)^2 p_i = D_x \quad (4.10)$$

Для безперервної випадкової величини:

$$m_2 = M[X^2] = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{-\infty}^{+\infty} (x - m_x)^2 f(x) dx = D_x \quad (4.11)$$

Другий центральний момент називається **дисперсією**. Дисперсія випадкової величини є характеристикою розсіювання цієї величини навколо математичного сподівання. У випадку, якщо це розсіювання відсутнє, величина D_x дорівнює нулю. Дисперсія має розмірність квадрата випадкової величини, що не завжди зручно. Тому як характеристику розсіювання часто використовують середнє квадратичне відхилення X:

$$\sigma_x = \sqrt{D_x}. \quad (4.12)$$

Центральні моменти більш високого порядку можуть характеризувати ступінь асиметрії розподілу випадкової величини, крутість кривої розподілу і т. інше.

Властивості числових характеристик

1. Математичне сподівання не випадкової величини дорівнює їй самій

$$M[c] = \sum_{i=1}^n c p_i = c \sum_{i=1}^n p_i = c. \quad (4.13)$$

2. Математичне сподівання добутку не випадкової величини C і випадкової величини X дорівнює добутку цієї не випадкової величини і математичного сподівання випадкової величини X :

$$M[cX] = \sum_{i=1}^n c x_i p_i = c \sum_{i=1}^n x_i p_i = cM[X], \quad (4.14)$$

тобто не випадкову величину C можна виносити за знак математичного сподівання.

3. Дисперсія не випадкової величини C дорівнює нулю:

$$D[c] = M[c^2] = M[(c - m_c)^2] = M[0] = 0. \quad (4.15)$$

4. Дисперсія добутку не випадкової величини C на випадкову величину X дорівнює добутку квадрата цієї не випадкової величини на дисперсію випадкової величини X :

$$D[cX] = M[c^2 X^2] = c^2 M[X^2] = c^2 D[X], \quad (4.16)$$

тобто не випадкову величину можна виносити з-під знака дисперсії, звівши її у квадрат.

**Тема 5. Найбільш важливі для практики закони розподілу
випадкових величин**

Біноміальний закон розподілу

Дискретна випадкова величина X має біноміальний закон розподілу (розподіл Бернуллі), якщо її можливі значення: $0, 1, \dots, n$, а відповідні ймовірності визначаються співвідношенням:

$$P_n(m) = C_n^m p^m q^{n-m} \quad (5.1)$$

де p - імовірність «успіху» в одному досліді, $0 < p < 1$; $q = 1 - p$.

Таким чином, розподіл залежить від двох параметрів n і p .

Для визначення числових характеристик випадкової величини, яка має біноміальний закон розподілу, введемо поняття **виробляючої функції**.

Нехай дискретна випадкова величина X приймає невід'ємні цілочисельні значення $0, 1, 2, \dots, k$ з ймовірностями p_1, p_2, \dots, p_k ($p_k = P\{X=k\}$).

Виробляючою функцією для випадкової величини X називається вираз виду:

$$\varphi(z) = \sum_{k=0}^{\infty} p_k z^k \quad (5.2)$$

де z – довільний параметр, $0 < z \leq 1$. Коефіцієнти p_k при z^k у виробляючій функції дорівнюють ймовірностям того, що випадкова величина X прийме значення k . Вираження (5.2) залишається справедливим і якщо число значень X скінченне, тому що при $k > n$ імовірності p_k обертаються в нуль.

При $z=1$ виробляюча функція дорівнює одиниці:

$$\varphi(1) = \sum_{k=0}^{\infty} p_k = 1. \quad (5.3)$$

Візьмемо похідну по z від виробляючої функції:

$$\varphi'(z) = \sum_{k=0}^{\infty} k p_k z^{k-1}$$

Нехай $z=1$, тоді $\varphi'(1) = \sum_{k=0}^{\infty} k p_k$, тобто похідна по z від виробляючої функції при $z=1$ являє собою суму добутків значень X на їхній імовірності, а отже є математичним сподіванням X .

$$M[X] = m_x = \varphi'(1). \quad (5.4)$$

Візьмемо другу похідну від $\varphi(z)$:

$$\varphi''(z) = \sum_{k=0}^{\infty} k(k-1)p_k z^{k-2}$$

Нехай $z=1$, одержимо $\varphi''(1) = \sum_{k=0}^{\infty} (k^2 - k)p_k = \sum_{k=0}^{\infty} k^2 p_k - \sum_{k=0}^{\infty} k p_k$. Перший доданок – другий початковий момент α_2 , а другий – математичне сподівання X . Дістанемо вираження другого початкового моменту α_2 .

$$\alpha_2 = \varphi''(1) + \varphi'(1). \quad (5.5)$$

Тобто другий початковий момент дорівнює сумі першої й другої похідних виробляючої функції при $z=1$.

Визначимо числові характеристики випадкової величини X , розподіленої за біноміальним законом. Запишемо виробляючу функцію:

$$\varphi(z) = \sum_{m=0}^n P_m z^m = \sum_{m=0}^n C_n^m p^m q^{n-m} z^m.$$

Такий же вигляд має n -я ступінь бінома:

$$\varphi(z) = (q + pz)^n. \quad (5.6)$$

Візьмемо похідну від вираження (5.6):

$$\varphi'(z) = np(q + pz)^{n-1}$$

і підставимо $z=1$, дістанемо математичне сподівання X :

$$m_x = \varphi'(1) = np(q + p)^{n-1} = np(1)^{n-1} = np. \quad (5.7)$$

Для обчислення дисперсії знайдемо другий початковий момент:

$$\begin{aligned} \alpha_2 &= \varphi''(1) + m_x \\ \varphi''(z) &= n(n-1)p^2(q + pz)^{n-2}; \quad \varphi''(1) = n(n-1)p^2 \\ \alpha_2 &= n(n-1)p^2 + np. \end{aligned}$$

$$D_x = \alpha_2 - m_x^2 = n(n-1)p^2 + np - n^2p^2 = n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p) = npq.$$

Таким чином, числові характеристики випадкової величини, розподіленої за біноміальним законом мають вигляд:

$$m_x = np; \quad D_x = npq; \quad \sigma_x = \sqrt{npq}. \quad (5.8)$$

Закон розподілу Пуассона

Можна показати, що при $n \rightarrow \infty$ границя вираження біноміального розподілу дорівнює

$$\lim_{n \rightarrow \infty} C_n^m p^m q^{n-m} = \frac{a^m}{m!} e^{-a}.$$

Отримана формула виражає закон розподілу Пуассона. Таким чином, коли імовірність p появи події A в кожному окремому досліді мала, а число дослідів n велике, біноміальний закон розподілу дискретної випадкової величини може бути приблизно замінений законом Пуассона:

$$P(m) = \frac{a^m}{m!} e^{-a} \quad (5.9)$$

Закон розподілу Пуассона визначається одним параметром $a = np$, що є одночасно математичним сподіванням і дисперсією випадкової величини X , розподіленої за законом Пуассона. Розподіл Пуассона з параметром $a = np$ можна приблизно застосовувати замість біноміального, коли число дослідів n дуже велике, а імовірність p дуже мала, тобто в кожному окремому досліді подія A з'являється вкрай рідко. Розподіл Пуассона часто використовується, коли ми маємо справу із числом подій, що з'являються на проміжку часу. Наприклад, число дефектів на новій ділянці шосе довжиною 10 км, число місць витoku води на 100 км водопроводу, число поломок надійного технічного устрою за певний період часу, наприклад, за рік.

Експонентний закон розподілу

Розглянемо потік однорідних подій, що ідуть одне за одним у випадкові моменти часу. Його можна розглядати як послідовність точок на числовій осі, що показана на рисунку 5.1.

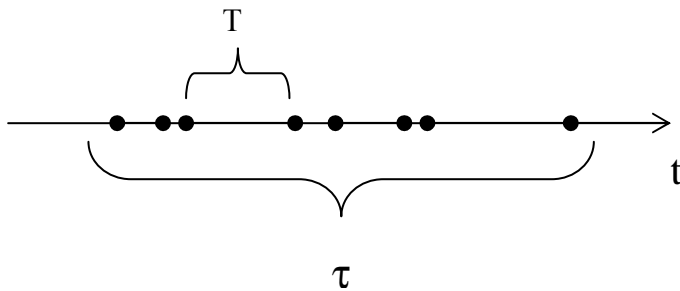


Рис.5.1 - Послідовність точок на числовій осі

Якщо середнє число подій λ на ділянці часу довжиною τ залежить тільки від довжини ділянки, тобто постійне, не залежить від числа подій, що потрапляють на інші ділянки, і події в потоці ідуть по одній, то такий потік називається **найпростішим**, або **стаціонарним пуассоновським потоком**.

Потік подій характеризується двома випадковими величинами: X – число подій, що потрапляють на інтервал часу τ і T – проміжок часу між двома випадковими подіями в потоці. Помітимо, що X є дискретною випадковою величиною, а T – безперервною.

Число подій, що потрапляють на ділянку часу τ у такому потоці має закон розподілу Пуассона (5.9), де $a = \lambda * \tau$ - математичне сподівання числа подій; λ - інтенсивність потоку подій (число подій в одиницю часу); τ - довжина ділянки часу.

Імовірність того, що на ділянці часу τ не відбудеться жодної події за законом Пуассона:

$$P(0) = \frac{a^0}{0!} e^{-a} = e^{-\lambda * \tau} \quad (5.10)$$

Отримана ймовірність стосовно випадкової величини проміжку часу між двома сусідніми подіями T , це імовірність того, що він виявиться більше деяко-

го поточного t : $P\{T < t\}$. Тоді імовірність того, що T виявиться менше деякого поточного t ($P\{T < t\}$), є за визначенням функцією розподілу T і визначиться як імовірність протилежної події:

$$F(t) = P\{T < t\} = 1 - P\{T \geq t\} = 1 - e^{-\lambda t}. \quad (5.11)$$

Щільність розподілу T визначимо як похідну функції розподілу $F(t)$:

$$f(t) = d(t)/dt = \lambda e^{-\lambda t}. \quad (5.12)$$

Такий закон розподілу називається **експонентним**.

Експонентний розподіл часто зустрічається в теорії масового обслуговування, а також у теорії надійності.

Визначимо числові характеристики випадкової величини, розподіленої за експонентним законом:

$$m_t = \int_{-\infty}^{+\infty} t f(t) dt = \int_0^{\infty} t \lambda e^{-\lambda t} dt = \lambda \int_0^{\infty} t e^{-\lambda t} dt$$

Проводячи інтегрування вроздріб, урахуємо, що при $t \rightarrow \infty$ $\exp\{-t\}$ прагне до нуля швидше, ніж зростає будь-який ступінь t , дістанемо:

$$m_x = 1/\lambda; \quad (5.13)$$

Таким чином, математичне сподівання випадкової величини, розподіленої за експонентним законом, зворотно параметру розподілу λ .

Визначимо дисперсію за формулою:

$$D_t = \alpha_2 - m_t^2 = \lambda \int_0^{\infty} t^2 e^{-\lambda t} dt - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}, \quad (5.14)$$

звідки середнє квадратичне відхилення:

$$\sigma_x = 1/\lambda, \quad (5.15)$$

де λ - параметр розподілу (середнє число подій в одиницю часу).

Знайдемо імовірність влучення випадкової величини, що має експонентний розподіл, в інтервал значень (α, β) . Нагадаємо, що ця ймовірність дорівнює приростові функції розподілу на інтервалі (α, β) . З формули (5.11) маємо $F(\alpha) = 1 - e^{-\alpha\lambda}$, $F(\beta) = 1 - e^{-\beta\lambda}$, тоді

$$P\{\alpha \leq t \leq \beta\} = F(\beta) - F(\alpha) = 1 - e^{-\lambda\beta} - 1 + e^{-\lambda\alpha} = e^{-\lambda\alpha} - e^{-\lambda\beta} \quad (5.16)$$

Нормальний закон розподілу ймовірностей

Цей закон також називається законом Гауса, оскільки був запропонований ним при дослідженні помилок точних вимірювань (помітимо, що помилки грубих вимірювань мають інший розподіл ймовірностей). Закон базується на двох посилках:

- 1) помилки різного знака, однакові по величині, мають однакову ймовірність;
- 2) малі помилки більш ймовірні, ніж великі (промахи).

Цим посилкам відповідає горбоподібна крива, симетрична щодо середнього значення помилки вимірювання (рис. 5.2.).

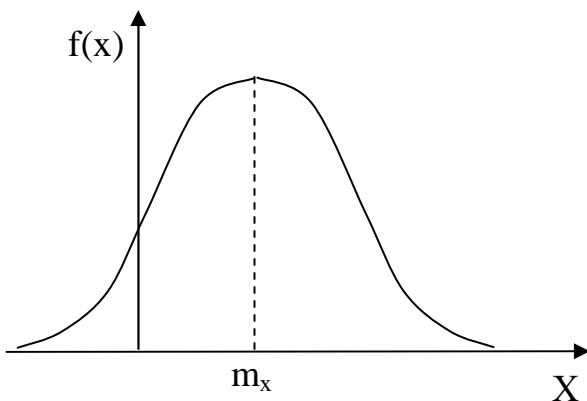


Рис.5.2 – Крива нормального розподілу

Отримана крива апроксимується вираженням:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \quad (5.17)$$

Звідси видно, що нормальний закон розподілу визначається двома параметрами:

На практиці закон нормального розподілу зустрічається дуже часто.

Тому що існує велике число нормально розподілених випадкових величин. Якщо відомі параметри m_x і σ_x , то із сімейства всіх кривих нормального розподілу виділяють одну з певною щільністю.

Ймовірність влучення нормально розподіленої випадкової величини в інтервал значень (α, β) дорівнює інтегралу від щільності розподілу в межах від α до β , але тому що інтеграл від $f(x)$ не береться в елементарних функціях, для

визначення ймовірностей, пов'язаних з нормально розподіленою випадковою величиною, користуються функцією Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \quad (5.18)$$

значення якої приведені в довідкових таблицях.

Імовірність влучення випадкової величини X на дільницю значень (α, β) виражається через функцію Лапласа формулою:

$$P\{\alpha < X < \beta\} = \Phi\left(\frac{\beta - m_x}{\sigma}\right) - \Phi\left(\frac{\alpha - m_x}{\sigma}\right) \quad (5.19)$$

При обчисленні ймовірностей користуються наступними властивостями функції Лапласа:

- 1) при $x=0$ $\Phi(x)=0$;
- 2) при $x=\infty$ $\Phi(x)=0,5$;
- 3) при $x=-\infty$ $\Phi(x)=-0,5$;
- 4) функція $\Phi(x)$ є непарною функцією, тобто $\Phi(-x) = -\Phi(x)$

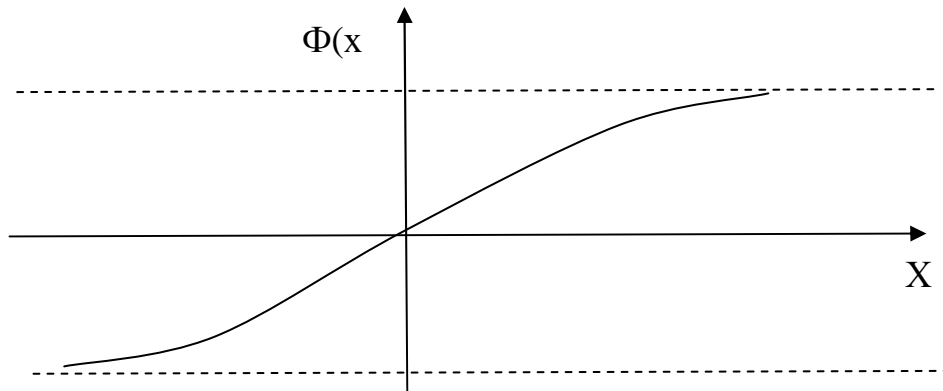


Рис. 5.3 – Крива функції Лапласа

Таким чином, всі можливі значення функції Лапласа належать інтервалу $(-0,5; +0,5)$, причому при $|x| > 4$ можна вважати, що $\Phi(x) \approx \pm 0,5$.

Скористаємося формулою (5.16), і визначимо функцію розподілу для випадкової величини, розподіленої нормально:

$$F(x) = \int_{-\infty}^x f(x) dx = \Phi\left(\frac{x - m_x}{\sigma_x}\right) - \Phi\left(\frac{-\infty - m_x}{\sigma_x}\right) = \Phi\left(\frac{x - m_x}{\sigma_x}\right) + 0,5. \quad (5.20)$$

Оскільки нормальний розподіл є симетричним, звичайно уявляє інтерес імовірність влучення нормально розподіленої випадкової величини в дільницю значень, симетричну щодо її математичного сподівання (рис 5.4):

$$P(|x - m_x| < l)$$

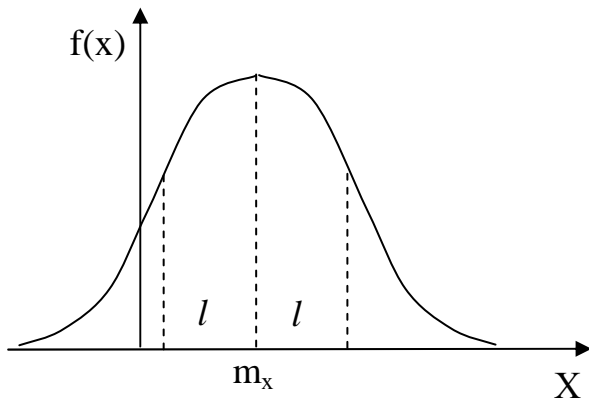


Рис.5.4 - Інтервал, симетричний щодо математичного сподівання

Для її визначення також скористаємося формулою (5.19).

$$\begin{aligned} P\{|x - m_x| < l\} &= P\{m_x - l < X < m_x + l\} = \\ &= \Phi\left(\frac{m_x + l - m_x}{\sigma_x}\right) - \Phi\left(\frac{m_x - l - m_x}{\sigma_x}\right) = 2\Phi\left(\frac{l}{\sigma_x}\right) \end{aligned}$$

Покладемо тепер $l = \sigma_x$, тоді:

$$P\{|x - m_x| < \sigma\} = 2 * \Phi\left(\frac{\sigma_x}{\sigma_x}\right) = 2 * \Phi(1) = 0,68$$

Таким чином, 68 % значень будь-якої нормально розподіленої випадкової величини лежать в інтервалі $(m_x \pm \sigma_x)$.

Нехай $l = 2\sigma_x$, тоді:

$$P\{|x - m_x| < 2\sigma\} = 2 * \Phi\left(\frac{2\sigma_x}{\sigma_x}\right) = 2 * \Phi(2) = 0,95.$$

Отже 95 % значень будь-якої нормально розподіленої випадкової величини лежать в інтервалі $(m_x \pm 2\sigma_x)$.

Якщо $l = 3\sigma_x$, то:

$$P\{|x - m_x| < 3\sigma\} = 2 * \Phi\left(\frac{3\sigma_x}{\sigma_x}\right) = 2 * \Phi(3) = 0,997.$$

Тобто 99,7 % значень будь-якої нормально розподіленої випадкової величини лежать в інтервалі $(m_x \pm 3\sigma_x)$.

Ця властивість називається «правилом трьох сигм».

Поняття про центральну граничну теорему

Центральна гранична теорема – це загальна назва групи теорем, що стосуються законів розподілу суми випадкових величин, зміст яких зводиться до наступного:

закон розподілу суми незалежних випадкових величин наближається до нормального закону розподілу при необмеженому збільшенні числа доданків, якщо всі величини мають кінцеві математичні сподівання й дисперсії й жодна з величин за значенням різко не відрізняється від інших.

Познайомимося із двома теоремами.

Теорема 1. Якщо незалежні X_1, X_2, \dots, X_n мають той самий закон розподілу з математичним сподіванням m і дисперсією D , то при необмеженому збільшенні n закон розподілу суми $X_1 + X_2 + \dots + X_n$ необмежено наближається до нормального.

Теорема 2 (центральна гранична теорема Ляпунова). Якщо випадкова величина $Y = X_1 + X_2 + \dots + X_n$, де вплив кожного з доданків на всю суму рівномірно малий, то величина Y має розподіл, близький до нормального, і тим ближче, чим більше n .

Дослід показує, що закон розподілу суми незалежних випадкових величин, порівняних за своїм розсіюванням, досить швидко наближається до нормального. Уже при числі доданків порядку десяти закон розподілу суми можна замінити нормальним. При цьому коштовно те, що закони розподілу випадкових величин, які додаються, можуть бути будь-якими, заздалегідь невідомими.

Багато випадкових величин можна розглядати як суму незалежних доданків. Наприклад, помилки вимірювань, відхилення розмірів деталей, число продаж деякого товару, обсяг прибутку від реалізації, валютні курси і т. інше.

Якщо є доданки X_i , що роблять переважний вплив на величину Y , то робити ствердження про нормальний розподіл Y не можна. У цьому випадку закон розподілу Y буде визначатися композицією законів розподілу доданків, вплив яких на Y великий.

Закон рівномірної щільності

Безперервна випадкова величина X має рівномірний розподіл на дільниці від α до β , якщо її щільність розподілу на цій дільниці постійна:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in (a, b); \\ 0 & \text{при } x \notin (a, b) \end{cases} \quad (5.21)$$

Наприклад, помилка X при грубому вимірюванні, яка може приймати з постійною щільністю імовірності будь-яке значення між двома сусідніми цілими діленнями. Грубий вимір відрізняється від точного в тім, що результат грубих вимірів при повторенні завжди той самий; при точному ж вимірі – результат від разу до разу змінюється. Іншим прикладом рівномірного розподілу є абсциса навмання поставленої точки на відрізку $[a, b]$.

Визначимо числові характеристики випадкової величини, розподіленої рівномірно.

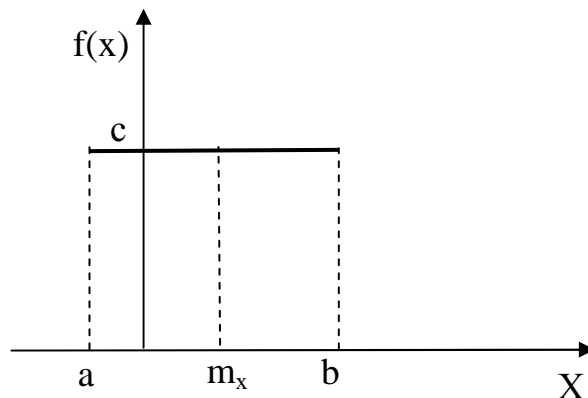


Рис.5.5 – Щільність рівномірного розподілу

Математичне сподівання:

$$m_x = \int_{-\infty}^{+\infty} xf(x)dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} * \frac{x^2}{2} \Big|_a^b = \frac{b+a}{2} \quad (5.22)$$

Дисперсія:

$$D_x = \int_{-\infty}^{+\infty} (x - m_x)^2 f(x)dx = \int_a^b (x - \frac{b+a}{2})^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12} \quad (5.23)$$

Середнє квадратичне відхилення:

$$\sigma_x = \sqrt{D_x} = \frac{b-a}{2\sqrt{3}} \quad (5.24)$$

Визначимо ймовірність влучення значень рівномірно розподіленої випадкової величини на інтервал (α, β) :

$$P\{\alpha < X < \beta\} = \int_{\alpha}^{\beta} \frac{1}{b-a} dx = \frac{\beta - \alpha}{b-a} \quad (5.25)$$

Функція рівномірного розподілу:

$$F(x) = \int_{\alpha}^x \frac{1}{b-a} dx = \frac{x - \alpha}{b-a}. \quad (5.26)$$

Тема 6. Система випадкових величин. Закони розподілу й числові характеристики системи

Багатомірна випадкова величина

При вивченні випадкових явищ іноді доводиться використовувати дві, три й більше випадкових величин. Спільний розгляд двох або декількох випадкових величин приводить до поняття системи випадкових величин. Систему декількох випадкових величин X, Y, \dots, W позначають (X, Y, \dots, W) і називають **багатомірною випадковою величиною**. При вивченні багатомірної випадкової величини недостатньо вивчити окремо випадкові величини, які її складають, а необхідно враховувати також зв'язки між цими величинами.

Найбільше практичне значення має система двох випадкових величин. Для характеристики системи двох випадкових величин використовують закони розподілу системи й числові характеристики системи.

Функція розподілу системи двох випадкових величин

Функція розподілу системи двох випадкових величин (X, Y) дорівнює імовірності того, що випадкова величина X прийме значення, менше x і випадкова величина Y прийме значення, менше y :

$$F(x, y) = P\{X \leq x, Y \leq y\} \quad (6.1)$$

Властивості функції розподілу системи:

4. Функція розподілу – неубутна функція, тобто

$$F(x_2, y) \geq F(x_1, y), \quad \text{якщо } x_2 > x_1 .$$

$$F(x, y_2) \geq F(x, y_1), \quad \text{якщо } y_2 > y_1 .$$

2. $F(-\infty, -\infty) = 0$.

3. $F(+\infty, +\infty) = 1$.

Щільність розподілу системи двох випадкових величин

Щільність розподілу ймовірностей системи двох випадкових величин (X, Y) $f(x, y)$ являє собою другу змішану похідну від $F(x, y)$:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \quad (6.2)$$

Тут припускається, що $F(x, y)$ безперервна й двічі диференційована. Властивості щільності розподілу ймовірностей:

1. Щільність розподілу невід'ємна, тобто $f(x, y) \geq 0$;
2. Подвійний інтеграл від щільності розподілу в нескінченних межах дорівнює одиниці:

$$\iint_{-\infty}^{+\infty} f(x, y) dx dy = 1.$$

Числові характеристики системи випадкових величин

Початковим моментом порядку k, s системи двох випадкових величин називається математичне сподівання добутку випадкових величин X, Y у ступені k і s :

$$\alpha_{k, s} = M[X^k Y^s] \quad (6.3)$$

Відповідно центральним моментом порядку k, s системи випадкових величин (X, Y) називається математичне сподівання добутку центрованих величин X, Y у ступені k і s :

$$\mu_{k, s} = M[X^k Y^s] \quad (6.4)$$

Для опису системи двох випадкових величин крім математичних сподівань і дисперсій X і Y використовують кореляційний момент і коефіцієнт кореляції. Кореляційним моментом є другий змішаний центральний момент:

$$\mu_{x, y} = M[X Y] = K_{xy} \quad (6.5)$$

Для дискретної випадкової величини K_{xy} визначається за формулою:

$$K_{x, y} = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_{ij} = \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)(y_j - m_y) p_{ij}, \quad (6.6)$$

де $p_{ij} = P\{X=x_i|Y=y_j\}$ – умовна ймовірність, тобто імовірність того, що X прийме значення x_i за умови, що Y прийме значення y_j .

Для безперервної випадкової величини:

$$K_{xy} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_x)(y - m_y) f(x, y) dx dy \quad (6.7)$$

Якщо події $P\{X=x_i\}$ і $P\{Y=y_j\}$ незалежні, то імовірність їхньої спільної появи за теоремою добутку дорівнює

$$p_{ij} = P\{X=x_i\} * P\{Y=y_j\} = p_i * p_j.$$

Тоді для кореляційного моменту справедливе вираження:

$$K_{x,y} = \sum_{i=1}^n x_i p_i \sum_{j=1}^m y_j p_j = 0,$$

тому що співмножники є центральними моментами першого порядку випадкових величин X і Y . Таким чином, кореляційний момент є характеристикою зв'язку між величинами X і Y , і у випадку незалежних X і Y він дорівнює нулю.

Як другий змішаний центральний момент кореляційний момент утримує також і розсіювання випадкових величин X і Y відносно один одного. Тому він не може характеризувати тісноту зв'язку між X і Y . Для визначення тісноти зв'язку між X і Y використовують коефіцієнт кореляції r_{xy} , який визначається за формулою:

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} \quad (6.8)$$

Переконаємося в тому, що r_{xy} характеризує ступінь тісноти лінійного зв'язку між двома випадковими величинами. Нехай випадкова величина Y функціонально (жорстко) залежить від випадкової величини X , причому залежність ця лінійна:

$$Y = a + bX$$

Визначимо математичне сподівання

$$M[Y] = M[a + bX] = \sum (ax_i + b)p_i = a \sum x_i p_i + \sum b p_i = a[X] + b.$$

Знайдемо дисперсію:

$$D[Y] = M[\overset{\circ}{Y}^2] = M[(Y-m_y)^2] = M[Y^2 - 2m_y Y + m_y^2] = \\ = M[(a+b)^2 - 2(am_x + b)(a+b) + (am_x + b)^2] = M[a^2(X-m_x)^2] = a^2 D_x,$$

а середнє квадратичне відхилення $\sigma_y = |a|\sigma_x$.

Визначимо коефіцієнт кореляції для жорстко зв'язаних X і Y , для чого виразимо $\overset{\circ}{Y}$ через $\overset{\circ}{X}$: $\overset{\circ}{Y} = Y - m_y = aX + b - am_x - b = a(X - m_x) = a\overset{\circ}{X}$, тоді кореляційний момент буде дорівнювати

$$K_{xy} = M[\overset{\circ}{Y}\overset{\circ}{X}] = [a\overset{\circ}{X}\overset{\circ}{X}] = aD_x$$

а коефіцієнт кореляції

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{aD_x}{\sigma_x |a|\sigma_x} = \frac{a\sigma_x^2}{\sigma_x |a|\sigma_x} = \frac{a}{|a|}.$$

Таким чином, якщо зв'язок функціональний, то коефіцієнт кореляції дорівнює 1, причому

$$r_{xy} = \begin{cases} -1 & \text{при } a < 0 \\ +1 & \text{при } a > 0 \end{cases} \quad (6.9)$$

У загальному випадку коефіцієнт кореляції лежить у межах $-1 \leq r_{xy} \leq +1$ і дорівнює нулю, якщо X і Y незалежні.

Поняття **корельованості** й **залежності** двох випадкових величин різні. Дві *корельовані* випадкові величини обов'язково залежні. Дві *залежні* випадкові величини необов'язково корельовані.

Функції випадкових величин

Функції однієї або декількох випадкових величин доводиться розглядати, коли аргументом деякої функції Y є система випадкових величин (X_1, X_2, \dots, X_n) , закон розподілу яких відомий. Функція Y є випадковою величиною, закон розподілу якої варто визначити. У більшості задач для визначення числових характеристик функції декількох випадкових величин досить знати тільки числові характеристики аргументів.

4. Математичне сподівання суми двох залежних або незалежних випадкових величин X і Y дорівнює сумі їхніх математичних сподівань:

$$M[X+Y] = \sum(x_i+y_i)p_i = \sum x_i p_i + \sum y_i p_i = M[X] + M[Y]. \quad (6.10)$$

Методом математичної індукції (узагальнення) дістанемо:

$$M[\sum X_i] = \sum M[X_i],$$

Математичне сподівання суми n випадкових величин дорівнює сумі їхніх математичних сподівань.

2. Математичне сподівання добутку двох випадкових величин X і Y дорівнює добутку їхніх математичних сподівань плюс кореляційний момент. Запишемо вираження для кореляційного моменту:

$$\begin{aligned} K_{xy} &= M[\overset{\circ}{Y}\overset{\circ}{X}] = M[(X - m_x)(Y - m_y)] = M[XY] - M[Xm_y] - M[Ym_x] + M[m_x m_y] = \\ &= M[XY] - m_x m_y - m_y m_x + m_x m_y = M[XY] - m_x m_y, \end{aligned}$$

звідки дістанемо

$$M[XY] = M[X] * M[Y] + K_{xy}. \quad (6.11)$$

Якщо випадкові величини X і Y незалежні, математичне сподівання добутку дорівнює добутку їхніх математичних сподівань.

3. Дисперсія суми двох випадкових величин X і Y дорівнює сумі дисперсій цих величин плюс подвоєний кореляційний момент:

$$\begin{aligned} D[X+Y] &= M[((X+Y)-M(X+Y))^2] = \\ &= M[X^2+2XY+Y^2 -2(X+Y)M(X+Y)+M^2(X+Y)] = \\ &= M[X^2+2XY+Y^2]-2M(X+Y)M(X+Y)+M^2(X+Y) = \\ &= M[X^2+2XY+Y^2]-(M[X]+ M[Y])^2 = \\ &= M[X^2]+ M[Y^2]-M^2[X]- 2M[X]*M[Y]- M^2[Y] = \\ &= D[X] + D[Y]+2M[XY] - 2M[X]M[Y] = D[X] + D[Y]+2K_{xy} \end{aligned}$$

Таким чином, дисперсія суми двох випадкових величин X і Y дорівнює сумі їхніх дисперсій плюс подвоєний кореляційний момент

$$D[X+Y] = D[X] + D[Y] + 2K_{xy}, \quad (6.12)$$

Якщо X і Y – незалежні випадкові величини, то дисперсія їхньої суми дорівнює сумі їхніх дисперсій, тоді сума n незалежних випадкових величин:

$$D[\Sigma X_i] = \Sigma D[X_i],$$

звідки середнє квадратичне відхилення суми:

$$\sigma_{\Sigma} = \sqrt{\Sigma \sigma_i^2}.$$

4. Дисперсія добутку двох незалежних випадкових величин X і Y визначається за формулою

$$D[XY] = D[X] * D[Y] + M[X]^2 D[Y] + M[Y]^2 D[X] \quad (6.13)$$

Тема 7. Закон великих чисел

Принцип практичної впевненості. Формулювання закону великих чисел

Якщо подія має дуже малу ймовірність (відмінну від нуля), то в одиничному випробуванні ця подія може наступити й не наступити. Але так міркуємо ми тільки теоретично, а на практиці вважаємо, що подія, яка має малу ймовірність, не наступить, і тому нехтуємо нею. Однак, у рамках математичної теорії не можна відповісти на запитання, якою повинна бути верхня границя ймовірності, щоб можна було назвати «практично неможливими» події, ймовірності яких не будуть перевищувати знайденої верхньої границі. Нехай, наприклад, робітник виготовляє на верстаті 100 виробів, з яких один в середньому виявляється бракованим. Очевидно, що ймовірність браку дорівнює 0,01, але нею можна зневажити й вважати робітника непоганим фахівцем. Але якщо будівельники будуть будувати будинки так, що з 100 будинків (у середньому) в одному будинку буде мати місце руйнування даху, то навряд чи можна зневажити ймовірністю такої події. Таким чином, у кожному окремому випадку варто виходити з того, наскільки важливі наслідки в результаті настання події.

Ймовірність, якою можна зневажити в даному дослідженні, називається рівнем значимості.

Сформулюємо принцип практичної впевненості: «Якщо випадкова подія має малу ймовірність (наприклад, $p < 0,01$), то при одиничному випробуванні можна практично вважати, що ця подія не відбудеться, а якщо подія має ймовірність, близьку до одиниці ($p > 0,99$), то практично при одиничному випробуванні можна вважати, що ця подія відбудеться напевно».

Основною закономірністю масових випадкових явищ є властивість усталеності середніх результатів. У широкому значенні слова під «законом великих чисел» розуміють відому із глибокої стародавності властивість усталеності масових випадкових явищ, що полягає в тому, що середній результат великої кі-

лькості випадкових явищ практично перестає бути випадковим і може бути передвіщений з достатньою визначеністю.

У вузькому значенні слова під «законом великих чисел» розуміють сукупність теорем, у яких встановлюється факт наближення середніх характеристик явища до деяких постійних величин у результаті великої кількості спостережень. Розглянемо деякі з них.

1. Лема Маркова. Якщо випадкова величина X не приймає від'ємних значень, то для будь-якого позитивного числа справедлива нерівність:

$$P\{X > \alpha\} \leq \frac{M[X]}{\alpha} \quad (7.1)$$

Доказ.

4. Нехай X — дискретна випадкова величина, задана рядом розподілу, причому $0 \leq x_1 < x_2 < \dots < x_n$.

X_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

Всі значення випадкової величини розіб'ємо на дві групи. До першої групи віднесемо значення, менші α (нехай це будуть x_1, x_2, \dots, x_k), до другої — всі інші значення більші або рівні α ($x_{k+1}, x_{k+2}, \dots, x_n$).

Як відомо, математичне сподівання дискретної випадкової величини визначається за формулою:

$$M[X] = x_1 p_1 + x_2 p_2 + \dots + x_k p_k + x_{k+1} p_{k+1} + \dots + x_n p_n.$$

Відкинемо в правій частині формули перші k доданків. Оскільки $p_i > 0$, $x_i \geq 0$, крім того, при $x_i \geq \alpha$, $i \geq k+1$, буде мати місце наступна нерівність: $M[X] \geq x_{k+1} p_{k+1} + \dots + x_n p_n \geq \alpha(p_{k+1} + \dots + p_n)$.

З того, що

$$p_{k+1} + \dots + p_n = P\{X=x_{k+1}\} + \dots + P\{X=x_n\} = P\{X \geq \alpha\},$$

слідкує, що:

$$M[X] \geq \alpha P\{X \geq \alpha\}.$$

Розділимо обидві частини останньої нерівності на α і дістанемо нерівність (7.1).

4. Нехай X – безперервна випадкова величина. Оскільки за умовою X не приймає невід’ємних значень, то її щільність імовірності $f(x) = 0$ при всіх $x < 0$. Тому

$$M[X] = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^{\infty} xf(x)dx \geq \int_{\alpha}^{\infty} xf(x)dx \geq \alpha \int_0^{\infty} xf(x)dx = \alpha P\{X \geq \alpha\},$$

розділимо на $\geq \alpha$, дістанемо нерівність (7.1). Лему доведено.

2. Нерівність Чебишева. Імовірність того, що відхилення випадкової величини X від її математичного сподівання за абсолютною величиною буде менше деякого додатного числа ε , обмежена знизу величиною

$$1 - \frac{D[X]}{\varepsilon^2} \quad \text{або} \quad P\{|X - M[X]| < \varepsilon\} \geq 1 - \frac{D[X]}{\varepsilon^2}. \quad (7.2)$$

Доказ.

Нехай маємо випадкову величину $(X - M[X])^2$. Оскільки вона приймає тільки додатні значення, до неї можна вжити лему Маркова, покладаючи в ній $\alpha = \varepsilon^2$, дістанемо:

$$P\{(X - M[X])^2 < \varepsilon^2\} \geq 1 - \frac{M[(X - M[X])^2]}{\varepsilon^2}.$$

Зазначимо, що $M[(X - M[X])^2] = D[X]$, а також урахуємо, що ймовірність $P\{(X - M[X])^2 < \varepsilon^2\}$ дорівнює ймовірності $P\{|X - M[X]| < \varepsilon\}$. На цій підставі можемо записати:

$$P\{|X - M[X]| < \varepsilon\} \geq 1 - \frac{D[X]}{\varepsilon^2}.$$

3. Теорема Чебишева. При необмеженому збільшенні числа n незалежних випробувань середня арифметична спостережуваних значень випадкової величини сходиться за імовірністю до її математичного сподівання, тобто для любого додатного ε

$$P\left\{\left|\bar{X} - M[X]\right| < \varepsilon\right\} = 1. \quad (7.3)$$

Ця теорема встановлює зв'язок між середньою арифметичною \bar{X} спостережуваних значень випадкової величини X і її математичним сподіванням $M[X]$.

Доказ.

Спостережувані значення X (x_1, x_2, \dots, x_n) у силу незалежності дослідів можна розглядати як незалежні випадкові величини, що мають однаковий розподіл (таке ж, як в X) з однаковими параметрами – математичним сподіванням $a = M[X]$ і дисперсією $D[X]$. За властивостями математичного сподівання й дисперсії можна записати:

$$\begin{aligned} M[\bar{X}] &= M\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n} M[X_1 + X_2 + \dots + X_n] = \frac{1}{n} (M[X_1] + M[X_2] + \dots + M[X_n]) = \\ &= \frac{1}{n} (a + a + \dots + a) = \frac{1}{n} na = a. \\ D[\bar{X}] &= D\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n^2} D[X_1 + X_2 + \dots + X_n] = \frac{1}{n^2} (D[X_1] + D[X_2] + \dots + D[X_n]) = \\ &= \frac{1}{n^2} nD[X] = \frac{D[X]}{n}. \end{aligned}$$

Скористаємося нерівністю Чебишева для випадкової величини X і підставимо в нього отримані параметри:

$$P\{|\bar{X} - a| < \varepsilon\} \geq 1 - \frac{D[X]}{n\varepsilon^2} \quad (7.4)$$

Якщо тепер в отриманій нерівності взяти як завгодно мале додатне ε і необмежено збільшити n , дістанемо

$$\lim_{n \rightarrow \infty} \left(1 - \frac{D[X]}{n\varepsilon^2}\right) = 1,$$

що й доводить теорему Чебишева.

З теореми Чебишева випливає важливий практичний висновок: невідоме значення математичного сподівання випадкової величини ми вправі замінити середнім арифметичним значенням, отриманим за досить великою кількістю дослідів. При цьому, чим більше проведено дослідів, тим з більшою ймовірністю можна чекати, що пов'язана із цією заміною помилка ($\bar{X} - a$) не перевер-

шить задану величину ϵ . При відомому значенні дисперсії $D[X]$, наприклад, за заданим значенням імовірності $P\{|\bar{X} - a| < \epsilon\}$ й максимальній припустимій помилці ϵ , визначити число необхідних дослідів n ; або по заданим P і n визначити ϵ ; а також, по заданим ϵ і n визначити границю ймовірності події $|\bar{X} - a| < \epsilon$.

4. Теорема Бернуллі. При необмеженому зростанні числа незалежних випробувань n відносна частота $\frac{m}{n}$ появи події A сходиться за імовірністю до його ймовірності P , тобто

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| \leq \epsilon\right) = 1, \quad (7.5)$$

За допомогою цієї теореми встановлюється зв'язок між відносною частотою події і її ймовірністю. Вона була доведена Я. Бернуллі (опублікована в 1713р.) і поклала початок теорії ймовірностей як науки.

Доказ.

Відносна частота $\frac{m}{n}$ є випадковою величиною, а отже характеризується математичним сподіванням і дисперсією:

$$M\left[\frac{m}{n}\right] = p; \quad D\left[\frac{m}{n}\right] = \frac{pq}{n}$$

Запишемо нерівність Чебишева для випадкової величини $\frac{m}{n}$:

$$P\left\{\left|\frac{m}{n} - p\right| < \epsilon\right\} \geq 1 - \frac{D\left[\frac{m}{n}\right]}{\epsilon^2}.$$

Остаточно дістанемо:

$$P\left\{\left|\frac{m}{n} - p\right| < \epsilon\right\} \geq 1 - \frac{pq}{n\epsilon^2}. \quad (7.6)$$

Як би малим не було число ϵ , при $n \rightarrow \infty$ величина дробі $\frac{pq}{n\epsilon^2} \rightarrow 0$, а

$$P\left\{\left|\frac{m}{n} - p\right| < \epsilon\right\} \rightarrow 1.$$

З теореми Бернуллі витікає, що при досить великій кількості випробувань відносна частота $\frac{m}{n}$ появи події практично втрачає свій випадковий характер, наближаючись до постійної величини P — імовірності даної події. У цьому й складається принцип практичної впевненості.

Незважаючи на те, що при необмеженому зростанні числа незалежних випробувань різниця $\left| \frac{m}{n} - p \right|$ може виявитися як завгодно малою, все ж таки не можна сказати, що $\lim_{n \rightarrow \infty} \frac{m}{n} = P$. Таке твердження було б невірним, тому що в даному питанні не виконуються необхідні умови, що входять до складу визначення поняття границі. Справді, може статися, що подія A буде відбуватися при всіх наступних випробуваннях, починаючи з деякого номера $n > N$ і тоді $\lim_{n \rightarrow \infty} \frac{m}{n} = P$, але не виключений і той випадок, коли починаючи з деякого номера $n > N$, подія A не буде відбуватися при жодному випробуванні, і тоді $\lim_{n \rightarrow \infty} \frac{m}{n} = 0$.

Таким чином, при необмеженому числі незалежних випробувань може статися, що $\frac{m}{n} \rightarrow p$, але цього може й не статися. Тоді виникає запитання про те, яка ж імовірність того, що $\frac{m}{n} \rightarrow p$? З теореми Бернуллі відповіді на це запитання не впливає, але в більш глибоких дослідженнях з теорії ймовірностей доводиться, що при $n \rightarrow \infty$ $P \left\{ \frac{m}{n} \rightarrow p \right\} = 1$. Отже, $\frac{m}{n} \rightarrow p$, не за типом границі, а за ймовірністю.

Змістовий модуль 3

ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Тема 8. Основні поняття. Суть вибіркового методу

Математична статистика розробляє методи опрацювання результатів спостережень масових випадкових явищ з метою виявлення наявних у них закономірностей. Збір статистичних даних провадять за спеціальними правилами статистичного спостереження.

Суть **вибіркового методу** полягає в тому, що висновки, зроблені на основі вивчення частини сукупності (випадкової вибірки), можна розповсюдити на всю сукупність (генеральну сукупність). Числові характеристики генеральної сукупності називаються генеральними параметрами. До них відносять математичне сподівання й дисперсію, які є параметрами розподілу досліджуваної ознаки. Їхні теоретичні значення ніколи не відомі, але їх можна оцінити за значеннями вибірових характеристик.

Оцінкою параметра називається числова характеристика, отримана в результаті обробки випадкової вибірки.

Статистичними аналогами параметрів розподілу є генеральна середня \bar{X} й генеральна дисперсія $\sigma_{ген}^2$. Розраховані за даними вибірки числові характеристики позначають \tilde{X} або $\bar{X}_{виб}$, $\sigma_{виб}$ або $S_{виб}$.

При здійсненні вибірки можливі помилки спостереження, до яких відносять помилки реєстрації й помилки репрезентативності. **Помилки реєстрації** виникають через неточності й погрішності при одержанні відомостей про одиниці сукупності, в результаті чого правдиве значення досліджуваної ознаки не збігається з його зареєстрованим значенням. **Помилки репрезентативності** являють собою різницю між вибіровими й генеральними характеристиками досліджуваної сукупності. Їх підрозділяють на систематичні й випадкові. **Систематичні помилки** пов'язані з тим, що структура вибірки відрізняється від структури генеральної сукупності. Звичайно це буває пов'язане з порушенням ви-

падковості відбору. **Випадкові помилки** репрезентативності пояснюються тим, що досліджується тільки частина сукупності. Таким чином, наявність випадкової помилки споконвічно властива вибірковому методу.

Введемо ряд визначень.

Об'єктом спостереження називається сукупність предметів або явищ, об'єднаних якою-небудь загальною ознакою або властивістю якісного або кількісного характеру. Усякий об'єкт статистичного спостереження складається з окремих елементів — **одиниць спостереження**. Для вивчення характерних властивостей об'єкта випадковим образом відбирають із всієї сукупності обмежене число одиниць спостереження.

Генеральною сукупністю називають об'єкт спостереження, що містить всю сукупність одиниць, з яких виробляється вибірка.

Вибірковою сукупністю, або просто **вибіркою**, називають сукупність одиниць спостереження, випадково відібраних з генеральної сукупності.

Розрізняють два типи випадкових вибірок: властиво випадкова повторна вибірка (схема повернутої кулі) і властиво випадкова безповторна вибірка (схема неповерненої кулі). Вибір схеми відбору залежить від характеру досліджуваного об'єкта.

Об'ємом сукупності (вибіркової або генеральної) називають число одиниць цієї сукупності. Число одиниць вибіркової сукупності позначають n , а генеральної — N .

У результаті статистичного спостереження одержують набір значень спостережуваної ознаки — дані. Значення ознаки при переході від одного елемента до іншого варіюють (змінюються).

Варіантою називають кожне окреме значення досліджуваної ознаки

$$x_1, x_2, \dots, x_n \dots$$

Частотою називають число, що показує, скільки разів зустрічається в сукупності та чи інша варіанта. Частоти позначають відповідно

$$m_1, m_2, \dots, m_n \dots$$

Первинною статистичною сукупністю називається неупорядкований набір значень досліджуваної ознаки, отриманих у результаті спостереження.

Першим кроком в обробці статистичних даних є впорядкування отриманих значень ознаки в порядку зростання або убутання, тобто побудова варіаційного ряду.

Варіаційним рядом називається таблиця, в одному рядку якої розташовуються варіанти x_1, x_2, \dots, x_n у зростаючому або зменшуваному порядку, а в іншій — відповідні ним частоти m_1, m_2, \dots, m_n ...

x_i	x_1	x_2	...	x_n
m_i	m_1	m_2	...	m_n

де $\sum_{i=1}^n m_i = n$.

Варіація досліджуваної ознаки може бути дискретною й безперервною.

Дискретною називається варіація, при якій окремі значення ознаки (варіанти) відрізняються один від другого на деяку скінченну величину (звичайно ціле число).

Безперервною називається варіація, при якій значення ознаки можуть відрізнитися одне від іншого на яку завгодно малу величину.

Замість абсолютних значень частот m_i звичайно використовують відносні — p_i^* . Для одержання відносних частот необхідно відповідну частоту розділити на суму всіх частот:

$$p_1^* = \frac{m_1}{\sum_{i=1}^n m_i}, \quad p_2^* = \frac{m_2}{\sum_{i=1}^n m_i}, \quad \dots, \quad p_n^* = \frac{m_n}{\sum_{i=1}^n m_i} \quad (8.1)$$

Сума всіх відносних частот дорівнює одиниці:

$$\sum_{i=1}^n p_i^* = 1 \quad (8.2)$$

Якщо варіація ознаки має безперервний характер, то весь діапазон отриманих значень розділяють на інтервали так, що кінець одного інтервалу є початком наступного. При цьому частоти належать не до окремого значення ознаки,

а до інтервалу значень. Звичайно як значення інтервалу приймають його середину.

У вибірці, що має дискретний характер варіації, число одиниць спостереження n може виявитися дуже великим. У результаті варіаційний ряд може виявитися громіздким і незручним для обробки. У такому випадку діапазон отриманих значень також поділяють на інтервали. Довжина i -го інтервалу позначається k_i і визначається за формулою:

$$k_i = x_{i\max} - x_{i\min} \quad (8.3)$$

де $x_{i\max}$ і $x_{i\min}$ – відповідно верхня й нижня границі i -го інтервалу.

При групуванні даних з метою побудови інтервального варіаційного ряду завжди встає питання про вибір оптимального числа інтервалів. Це пов'язане з тим, що занадто велике число інтервалів зробить варіаційний ряд громіздким, а занадто мале приведе до огрубіння результатів статистичного аналізу. Для визначення оптимальної довжини інтервалу рекомендується використовувати формулу Старджеса:

$$k = \frac{x_{\max} - x_{\min}}{1 + 3,322 \lg n} \quad (8.4)$$

де x_{\max} і x_{\min} – відповідно найбільше й найменше значення варіантів ряду; n – число одиниць сукупності.

Для наочності варіаційний ряд можна зобразити графічно у вигляді полігона розподілу, кумулятивної кривої, або гістограми.

Полігон розподілу (дослівно — багатокутник розподілу) будується в прямокутній системі координат. Величина ознаки відкладається на осі абсцис, частоти або відносні частоти – по осі ординат (рис. 8.1).

Кумулятивна крива (кумулята) утворюється при зображенні варіаційного ряду з накопиченими відносними частотами в прямокутній системі координат (рис. 8.2). Накопичена частота певної варіанти утворюється підсумовуванням всіх частот варіант, що передують даній, із частотою цієї варіанти. При побудові кумуляти дискретної ознаки по осі абсцис відкладають значення ознаки (варіанти). Ординатами служать вертикальні відрізки, довжина яких пропор-

ційна накопиченій відносній частоті тієї або іншої варіанти Σp_i^* . З'єднавши вершини ординат прямими лініями, одержимо ламану (криву) кумуляту. При $\Delta x_i \rightarrow 0$ кумулята прагне до безперервної кривої, що представляє собою функцію розподілу випадкової величини.

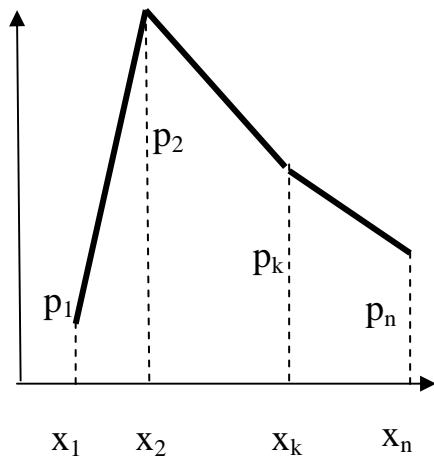


Рис. 8.1 – Полігон розподілу

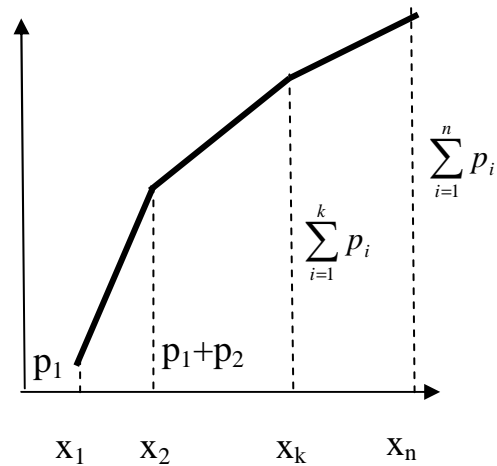


Рис. 8.2 – Кумулятивна крива

Гістограма розподілу будується аналогічно полігону в прямокутній системі координат (рис. 8.3). При побудові гістограми на осі абсцис вибирають відрізки, що відповідають інтервалам, на яких будують прямокутники із площею, пропорційною відносним частотам інтервалів p_i^* . З умови побудови гістограми випливає, що вся її площа дорівнює 1. Очевидно також, що при $\Delta x_i \rightarrow 0$ гістограма прагне до безперервної залежності, що уявляє собою щільність розподілу випадкової величини.

Визначення закону розподілу спостережуваної ознаки за статистичними даними

Однією із задач математичної статистики є одержання оцінок числових характеристик випадкових величин. Іншою задачею є визначення закону розподілу випадкової величини за статистичним даними. Оскільки ці дані завжди обмежені, то виникає задача згладжування або вирівнювання статистичних рядів за допомогою аналітичних виразів, що є найбільш компактним вираженням закономірності. Статистичний (варіаційний) ряд дозволяє побудувати шукані статистичні закони розподілу.

Зокрема, для того, щоб знайти статистичну функцію розподілу $F^*(x)$, досить підрахувати число варіантів, у яких ознака X прийняла значення менші x , тобто $X < x$. Якщо число таких варіантів $m(x)$, а обсяг сукупності дорівнює n , то

$$F^*(x) = \frac{m(x)}{n} \quad (8.5)$$

Якщо результати спостереження зведені в згрупований статистичний ряд, то на його підставі будують гістограму, з урахуванням того, що частота появ ознаки на i -м інтервалі буде

$$p_i^* = \frac{m_i}{n},$$

де m_i – кількість появ x на i -м інтервалі.

Для оформлення статистичного ряду у вигляді гістограми по осі абсцис відкладають інтервали ряду, а потім на кожному інтервалі будують прямокутник, площа якого дорівнює p_i^* .

Задача вирівнювання статистичних рядів зводиться до підбору теоретичної кривої розподілу, що виражає лише істотні риси статистичного матеріалу. Якщо клас функцій, що описують розподіл, відомий, то задача зводиться до раціонального вибору параметрів розподілу.

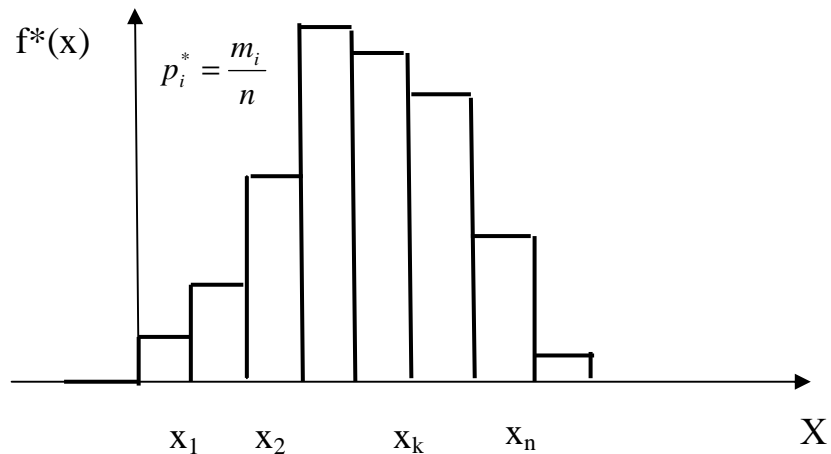


Рис. 8.3. Гістограма

Наприклад, гістограма, наведена на рис. 8.3, наводить на думку про нормальний розподіл ознаки X :

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

Тоді задача зводиться до відшукування двох параметрів – вибіркової середньої \bar{x} і вибіркового середнього квадратичного відхилення s .

Одним з методів рішення поставленої задачі є метод моментів (початкових і центральних), що був запропонований в 1894 році англійцем Пірсоном. Відповідно до цього методу числові параметри розподілу вибираються з таким розрахунком, щоб кілька найважливіших числових характеристик (моментів) були рівні їхнім статистичним оцінкам. Зокрема, математичне сподівання й дисперсія приймаються рівними їхнім статистичним оцінкам – вибіркової середньої й вибіркової дисперсії:

$$m_x = \tilde{X}, \quad D_x = \sigma_{\text{виб}}^2.$$

Іншим методом визначення статистичних оцінок параметрів розподілу є метод максимальної правдоподібності, запропонований в 1912 році англійцем Фішером.

Тема 9. Числові характеристики варіаційного ряду

Однією з найважливіших характеристик варіаційного ряду є його середнє значення. Розрізняють середню арифметичну зважену й незважену. Середня арифметична зважена визначається як відношення суми добутоків варіантів на відповідні їм частоти до суми всіх частот:

$$\bar{x} = \frac{\sum_{i=1}^k x_i m_i}{\sum_{i=1}^k m_i} \quad \text{або} \quad \bar{x} = \sum_{i=1}^k x_i p_i^* \quad (9.1)$$

де m_i – частоти варіаційного ряду; p_i^* – відносні частоти; k – число груп з однаковими значеннями ознаки.

Для розрахунку незваженої середньої арифметичної використовується формула

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (9.2)$$

де n – число елементів варіаційного ряду.

Залежно від особливостей досліджуваного явища крім арифметичної використовують геометричну, гармонійну, квадратичну, кубічну й інші середні величини.

Розглянемо ряд характеристик, які використовуються для вимірювання варіації ознаки.

Варіаційний розмах R , або широта розподілу – різниця між найбільшим і найменшим значеннями варіаційного ряду:

$$R = x_{\max} - x_{\min} \quad (9.3)$$

Варіаційний розмах застосовується для приблизної оцінки варіації, тому що він говорить тільки про відстань між найбільшим і найменшим значеннями варіантів варіаційного ряду. Більш точну міру варіації визначають за допомогою середнього лінійного відхилення, дисперсії й стандартного відхилення (або середнього квадратичного відхилення).

Середнє лінійне відхилення, або просте середнє відхилення (позначається d) являє собою середню арифметичну абсолютних значень відхилень варіант від середньої. Обчислюють середнє лінійне відхилення незважене або зважене:

$$d = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \qquad d = \frac{\sum_{i=1}^k |x_i - \bar{x}| m_i}{\sum_{i=1}^k m_i} \qquad (9.4)$$

Дисперсія варіаційного ряду – це середня арифметична квадрата відхилення значень ознак ряду від їхньої середньої арифметичної. Дисперсія незважена і зважена обчислюється за формулами:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \qquad \sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 m_i}{\sum_{i=1}^k m_i} \qquad (9.5)$$

Стандартне відхилення варіаційного ряду визначається як арифметичне значення квадратного кореня з дисперсії

$$\sigma = \sqrt{\sigma^2} \qquad (9.6)$$

Дисперсія – це міра розсіювання варіантів щодо середньої арифметичної. Чим більше варіація, тим далі від середньої знаходяться можливі значення ознак.

Коефіцієнт варіації V являє собою долю стандартного відхилення від середнього значення варіаційного ряду:

$$V = \frac{\sigma}{\bar{x}} \qquad (9.7)$$

Таким чином, коефіцієнт варіації є відносною мірою варіації, в той час як стандартне відхилення – абсолютна міра розсіювання варіантів ряду. Чим менше значення коефіцієнта варіації, тим однорідніше сукупність за досліджуваною ознакою.

Властивості вибірових числових характеристик

Будь-які значення вибірових характеристик, обчислені на підставі обмеженого числа елементів вибірки, містять елемент випадковості. Очевидно, що обробка декількох вибірок однакового об'єму дасть ряд різних оцінок відповідної числової характеристики. Отже, оцінки числових характеристик є випадковими величинами на відміну від самих числових характеристик, значення яких не випадкові. Необхідно, щоб помилка від заміни правдивого значення числової характеристики його наближеною оцінкою була мінімальною. Такій вимозі задовольняють оцінки числових характеристик, що мають властивості спроможності, незміщеності й ефективності.

Оцінка параметра a^* є **спроможною**, якщо при $n \rightarrow \infty$ вона сходиться за імовірністю до оцінюваного параметра a :

$$\lim_{n \rightarrow \infty} a^* = a. \quad (9.8)$$

Зокрема, на підставі теореми Чебишева вибірова середня \tilde{X} , визначена за формулами (9.1) і (9.2), є спроможною оцінкою генеральної середньої \bar{X} . Формули (9.5) дають спроможну оцінку генеральної дисперсії.

Оцінка параметра a^* є **незміщеною** (тобто не містить систематичної помилки), якщо її математичне сподівання дорівнює оцінюваному параметру a :

$$M[a^*] = a. \quad (9.9)$$

Вибіркова середня \tilde{X} , визначена за формулами (9.1) і (9.2), є лінійною функцією n незалежних випадкових величин x_i , тому вона сама є випадковою величиною, а отже має свої числові характеристики: математичне сподівання й дисперсію. Покажемо, що математичне сподівання вибірової середньої не залежить від числа дослідів n і дорівнює генеральній середній:

$$M[\tilde{X}] = M\left[\frac{\sum x_i}{n}\right] = \frac{1}{n} M\left[\sum x_i\right] = \frac{1}{n} * n * M[X] = \bar{X}.$$

Визначимо дисперсію оцінки вибірової середньої:

$$\sigma^2[\tilde{X}] = \sigma^2 \left[\frac{\sum x_i}{n} \right] = \left(\frac{1}{n} \right)^2 (\sum \sigma_i^2) = \left(\frac{1}{n} \right)^2 n \sigma^2 = \frac{\sigma_{ген}^2}{n}, \quad (9.10)$$

звідки одержимо середнє квадратичне відхилення вибіркової середньої:

$$\sigma[\tilde{X}] = \frac{\sigma_{ген}}{\sqrt{n}}. \quad (9.11)$$

Очевидно, що зі збільшенням числа дослідів n $\sigma^2[\tilde{X}]$ прагне до нуля, що свідчить про прагнення \tilde{X} до не випадкової величини \bar{X} .

Таким чином, вибіркова середня є незміщеною оцінкою генеральної середньої.

Розглянемо вибірку дисперсію. Можна показати, що математичне сподівання вибіркової дисперсії, визначене за формулою

$$\sigma_{выб}^2 = \frac{\sum (x_i - \tilde{X})^2}{n} \quad (9.12)$$

не дорівнює генеральній дисперсії, тобто є зміщеною оцінкою:

$$M[\sigma_{выб}^2] = \frac{n-1}{n} \sigma_{ген}^2$$

Користуючись цією оцінкою, ми будемо робити систематичну помилку в меншу сторону. Щоб її позбутися, варто ввести виправлення – помножити оцінку дисперсії, отриману за формулою (9.12), на $n/(n-1)$. Незміщена оцінка дисперсії називається також виправленою дисперсією. Вона позначається S^2 і визначається за формулою:

$$S^2 = \frac{n}{n-1} \sigma_{выб}^2 = \frac{\sum (x_i - \tilde{X})^2}{n-1} \quad (9.13)$$

Можна показати, що S^2 є незміщеною оцінкою генеральної дисперсії:

$$M[S^2] = \frac{n}{n-1} M[\sigma_{выб}^2] = \frac{n}{n-1} * \frac{n-1}{n} * \sigma_{ген}^2.$$

Якщо число спостережень n велике, то значення вибіркової дисперсії, обчисленої за формулою (9.12) практично збігається зі значенням, обчисленим за формулою (9.13). При $n > 50$ вони практично не відрізняються.

Оцінка параметра a^* є **ефективною**, якщо при заданому об'ємі вибірки вона має найменшу дисперсію. Ступінь ефективності оцінюють відношенням дисперсій:

$$F = \frac{\sigma_1^2}{\sigma_2^2}, \quad (9.14)$$

тобто якщо $F > 1$, то σ_2^2 більш ефективна, і навпаки.

Якщо на підставі статистичних даних необхідно визначити оцінки числових характеристик системи двох випадкових величин X і Y , то крім вибірових середніх і вибірових дисперсій знаходять оцінку кореляційного моменту за формулою:

$$K_{xy}^* = \frac{\sum (X_i - m_x^*)(Y_i - m_y^*)}{n - 1} \quad (9.15)$$

Наведені оцінки є спроможними й незміщеними.

Довірчий інтервал і довірна імовірність

Розглянуті оцінки параметрів є точковими, тому що характеризуються одним числом. Якщо точкова оцінка параметра визначена на підставі вибірки малого об'єму, вона може значно відрізнитися від оцінюваного параметра. Для визначення помилки від заміни генерального параметра його оцінкою використовують поняття довірчого інтервалу й довірчої імовірності.

Нехай для деякого параметра розподілу, наприклад, математичного сподівання m_x отримані спроможна й незміщена оцінка a^* . Потрібно визначити можливу помилку l . Призначимо досить велику імовірність β (0,95; 0,99) таку, що подію $A = \{ |a^* - m_x| < l \}$, яка характеризується цією ймовірністю, можна вважати практично достовірною. Знайдемо значення l , для якого справедлива рівність

$$P(A) = P\{(a - l) < m_x < (a + l)\} = \beta. \quad (9.16)$$

Рівність (9.16) означає, що діапазон можливої помилки при заміні m_x на a^* з імовірністю β буде дорівнювати $\pm l$. Тут a^* і β відомі, l підлягає визначен-

ню. З імовірністю β невідоме значення параметра m_x буде знаходитися в інтервалі $L=[a^*- 1, a^* + 1]$. Більші, ніж 1 за абсолютним значенням помилки, будуть зустрічатися з імовірністю

$$\alpha = 1 - \beta.$$

Зауважимо, що величина m_x не випадкова, а випадковим є положення відрізка $L=2*1$ на осі абсцис, обумовлене випадковою величиною a^* . Таким чином, довірчу ймовірність β можна трактувати як імовірність того, що випадковий інтервал L накріє точку m_x (рис. 9.1).

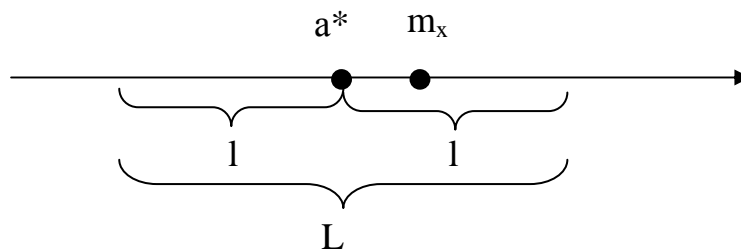


Рис. 9.1 – Довірчий інтервал.

Таким чином, значення β і L характеризують ступінь упевненості й величину погрішності при визначенні правдивого значення шуканого параметра m_x .

Приклад. Нехай зроблено n незалежних дослідів, з яких визначені значення вибіркової середньої \tilde{X} й вибіркової дисперсії S^2 . Потрібно визначити довірчий інтервал для вибіркової середньої.

Оскільки вибіркова середня визначається за формулою

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n},$$

її можна розглядати як функцію n незалежних випадкових величин x_i , які в загальному випадку можуть бути підлеглі будь-яким законам розподілу. Тоді, відповідно до центральної граничної теореми, закон розподілу вибіркової середньої \tilde{X} буде близький до нормального з параметрами:

$$M[\tilde{X}] = \bar{x};$$

$$D[\tilde{X}] = D\left[\frac{\sum_{i=1}^n x_i}{n}\right] = \frac{1}{n^2} \sum D[x_i] = \frac{nD[\tilde{X}]}{n^2} = \frac{D[\tilde{X}]}{n}$$

$$\sigma_{\text{внб}} = \frac{\sigma[\tilde{X}]}{\sqrt{n}}$$

Припустимо, що генеральна дисперсія відома, тоді скориставшись інтегралом імовірності

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

і врахувавши, що помилка симетрична відносно a^* , а β задана, дістанемо:

$$P\{(a^* - l) < \tilde{X} < (a^* + l)\} = \beta = \Phi\left(\frac{l}{\sigma[\tilde{X}]}\right).$$

Користуючись таблицею значень інтеграла ймовірностей, можемо знайти аргумент

$$\frac{l}{\sigma[\tilde{X}]} = \Phi^{-1}(\beta), \text{ звідси } l = \sigma[\tilde{X}] \Phi^{-1}(\beta).$$

Довірчий інтервал $L = 2 \cdot l$.

Тема 10. Елементи теорії кореляції

Метою кореляційного аналізу є визначення форми залежності між випадковими величинами X і Y і оцінка тісноти зв'язку між ними.

Залежність

$$y = f(x), \tag{10.1}$$

в якій кожному значенню X відповідає одне певне значення Y , називається **функціональною**.

Одному значенню випадкової величини X x_i може відповідати ряд значень Y : y_1, y_2, \dots, y_k , що може бути викликано впливом різних факторів на випадкову величину Y або помилками виміру. У цьому випадку залежність називається

вається **статистичною**. Для кожного значення x_i можна визначити умовне середнє \bar{y}_i (рис. 10.1).

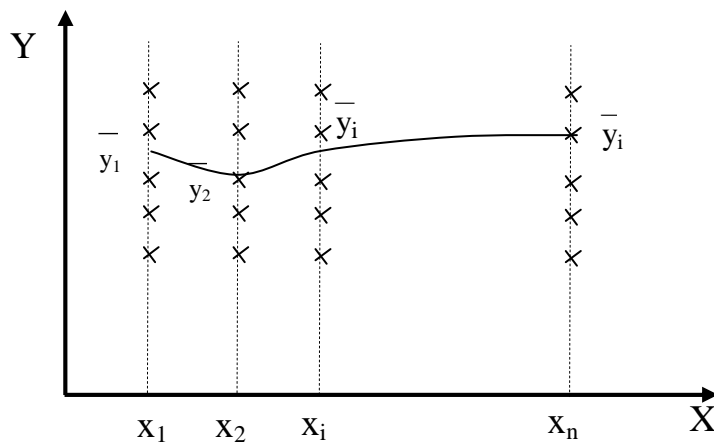


Рис. 10.1 – Статистична залежність.

Статистичною називається залежність між X і Y , при якій зі зміною випадкової величини X змінюється розподіл випадкової величини Y . Якщо при зміні X змінюється середнє значення Y , то така статистична залежність називається **кореляційною**.

$$\bar{y}_x = \varphi(x) \quad (10.2)$$

Кореляційний аналіз заснований на використанні рівняння регресії.

Регресією Y на X називається умовне математичне сподівання випадкової величини Y за умови, що X прийняла значення x_i . Лінія, що з'єднує точки \bar{y}_i , називається **лінією регресії**. Для апроксимації лінії регресії аналітичним вираженням використовують **рівняння регресії**. На практиці найчастіше використовують лінійне рівняння регресії:

$$Y = \rho_{yx} x + b \quad (10.3)$$

Коефіцієнт при x ρ_{yx} називається коефіцієнтом регресії.

Метод найменших квадратів

Для визначення значень параметрів ρ_{yx} і b рівняння регресії (10.3) застосовується *метод найменших квадратів* (МНК), що дозволяє при відомому класі залежності $\bar{y}_x = \varphi(x)$ так вибрати їхні значення, щоб вона щонайкраще відображала дані спостережень.

При використанні МНК вимога найкращого узгодження $\bar{y}_x = \varphi(x)$ з дослідними даними зводиться до того, щоб сума квадратів відхилень кривої, що згладжує залежність, від експериментальних точок оберталася в мінімум:

$$\sum_{i=1}^n (y_{ip} - y_i)^2 \rightarrow \min . \quad (10.4)$$

де y_i – значення Y , отримані в результаті спостережень;

y_{ip} – розрахункові значення Y , отримані за вираженням кривої, що згладжує $\varphi(x)$.

Якщо всі виміри провадилися з однаковою точністю й помилки вимірів розподілені за нормальним законом, то знайдена залежність буде найбільш ймовірною із всіх можливих у даному класі функцій.

З огляду на те, що $y_{ip} = \varphi(x_i)$, вираження (9.4) можна записати у вигляді

$$\sum_{i=1}^n [\varphi(x_i) - y_i]^2 \rightarrow \min \quad (10.5)$$

Невідомі параметри шуканої залежності визначають, записавши її не тільки як функцію аргументу x , але і як функцію невідомих параметрів a_j .

$$\sum_{i=1}^n [\varphi(x_i, a_1, a_2, \dots, a_j, \dots, a_m) - y_i]^2 \rightarrow \min \quad (10.6)$$

Умова (10.6) виконується, якщо всі часткові похідні суми квадратів відхилень за параметрами a_j будуть дорівнювати нулю. Часткові похідні дають систему $m+1$ рівнянь із $m+1$ невідомими, розв'язання якої дає шукані параметри a_j , що задовольняють умові (10.5).

Дістанемо для лінійного рівняння регресії (10.3) методом найменших квадратів вираження для коефіцієнта регресії ρ_{yx} і вільного члена b . Для цього підставимо в (10.6) вираження (10.3)

$$\sum_{i=1}^n [\rho_{yx} x_i + b - y_i]^2 \rightarrow \min$$

Для відшукання мінімуму візьмемо похідні за параметрами ρ_{yx} і b і дорівняємо їх до нуля, дістанемо систему рівнянь:

$$\left. \begin{aligned} 2 \sum_{i=1}^n [\rho_{yx} x_i + b - y_i] * x_i &= 0 \\ 2 \sum_{i=1}^n [\rho_{yx} x_i + b - y_i] &= 0 \end{aligned} \right\}, \quad (10.7)$$

з якої в результаті перетворень отримаємо:

$$\left. \begin{aligned} \rho_{yx} * \sum x_i^2 + b * \sum x_i &= \sum x_i y_i \\ \rho_{yx} * \sum x_i + nb &= \sum y_i \end{aligned} \right\} \quad (10.8)$$

звідки виразимо ρ_{yx} і b

$$\begin{aligned} \rho_{yx} &= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \\ b &= \frac{\sum x_i^2 * \sum y_i - \sum x_i * \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} \end{aligned} \quad (10.9)$$

Вибірковий коефіцієнт кореляції

Для оцінки тісноти лінійної кореляційної залежності служить вибірковий коефіцієнт кореляції. Для його визначення підставимо у вираження (9.8) наступні співвідношення:

$$\bar{x} = \frac{\sum x_i}{n}, \text{ звідки } \sum x_i = \bar{x}n;$$

$$\bar{y} = \frac{\sum y_i}{n}, \text{ звідки } \sum y_i = \bar{y}n;$$

$$\overline{x^2} = \frac{\sum x_i^2}{n}, \text{ звідки } \sum x_i^2 = \overline{x^2}n;$$

Одержимо

$$\left. \begin{aligned} \rho_{yx} * \overline{x^2}n + b * \bar{x}n &= \sum x_i y_i \\ \rho_{yx} * \bar{x} + b &= \bar{y} \end{aligned} \right\} \quad (10.10)$$

із другого рівняння виразимо b

$$b = \bar{y} - \rho_{yx} * \bar{x} \quad (10.11)$$

і підставивши його в перше рівняння, визначимо коефіцієнт регресії

$$\begin{aligned} \rho * \sum x_i^2 + (\bar{y} - \rho * \bar{x}) \sum x_i &= \sum x_i y_i \\ \rho * (\sum x_i^2 - \bar{x} * \sum x_i) &= \sum x_i y_i - \bar{y} * \sum x_i \\ \rho &= \frac{\sum x_i y_i - \bar{y} * \sum x_i}{\sum x_i^2 - \bar{x} * \sum x_i} = \frac{\sum x_i y_i - n \bar{y} \bar{x}}{n \overline{x^2} - n \bar{x}^2} = \frac{\sum x_i y_i - n \bar{y} \bar{x}}{n * \sigma_x^2}. \end{aligned} \quad (10.12)$$

Помножимо отриману рівність на дріб $\frac{\sigma_x}{\sigma_y}$ і одержимо

$$\rho * \frac{\sigma_x}{\sigma_y} = \frac{\sum x_i y_i - n \bar{y} \bar{x}}{n * \sigma_x \sigma_y},$$

де $\rho * \frac{\sigma_x}{\sigma_y} = r_g$ - вибірковий коефіцієнт кореляції.

Підставивши в рівняння $\bar{y}_x = \rho_{yx} x + b$ вираження для b (10.11), дістанемо рівняння регресії в наступному виді:

$$\bar{y}_x - \bar{y} = \rho_{xy} * (x - \bar{x}) \quad \text{або} \quad \bar{y}_x - \bar{y} = r_g \frac{\sigma_y}{\sigma_x} * (x - \bar{x}) \quad (10.13)$$

Коефіцієнт кореляції r_g має важливе значення. Він дозволяє оцінити величину лінійного зв'язку між двома випадковими величинами X і Y . Покладемо в рівнянні (10.13) $r_g = 0$, тоді

$$\bar{y}_x - \bar{y} = 0$$

або

$$\bar{y}_x = \bar{y},$$

тобто при $r_g = 0$ всі умовні середні рівні вибіркового середньому, тобто при зміні випадкової величини X випадкова величина Y не змінюється й графік рівняння регресії паралельний осі абсцис. Це говорить про те, що Y не залежить від X , між ними немає лінійного зв'язку. Однак X і Y можуть бути зв'язані нелінійним зв'язком, який може виявитися як кореляційним так і функціональним.

Дисперсія Y у точці $X=x_i$ щодо умовного середнього S_y визначається за формулою

$$S_y = D_y(1 - r_g^2), \quad (10.14)$$

де D_y – дисперсія Y щодо загального середнього.

Покладемо в цій формулі $r_g = 1$, тоді

$$S_y = 0,$$

тобто розсіювання значень Y у кожній точці відсутнє, рівняння (10.13) буде мати вигляд $\bar{y}_x - \bar{y} - \frac{\sigma_y}{\sigma_x} * (x - \bar{x}) = 0$, тобто будь-яка пара чисел x і y йому задовольняє. Звідси витікає, що при $r_g = 1$ між X і Y існує функціональний лінійний зв'язок.

З формули (10.14) також витікає, що зі збільшенням r_g дисперсія відносно умовного середнього S_y зменшується, тобто зменшується розсіювання навколо умовних середніх, а значить тіснота зв'язку збільшується.

Таким чином, вибіркового коефіцієнта кореляції приймає значення від -1 до +1, і характеризує тісноту лінійного зв'язку між ознаками у вибірці. Якщо $r_B = 0$, то лінійний зв'язок відсутній, чим ближче значення $|r_B|$ до одиниці, тим тісніше зв'язок, і при $r_B = 1$ він стає функціональним.

Вибіркове кореляційне відношення

Для оцінки тісноти нелінійного кореляційного зв'язку застосовують вибіркового кореляційного відношення η . Вибірковим **кореляційним відношенням** Y до X називається відношення міжгрупового середнього квадратичного відхилення до загального середнього квадратичного відхилення ознаки Y :

$$\eta_{yx} = \frac{\sigma_{y_x}^-}{\sigma_y} \quad (10.15)$$

де $\sigma_{y_x}^-$ - міжгрупове середнє квадратичне відхилення, визначається як квадратний корінь із міжгрупової дисперсії за формулою

$$\sigma_{y_x}^- = \sqrt{D_{\text{межгр}}} = \sqrt{\frac{\sum N_j (\bar{y}_j - \bar{y})^2}{n}}, \quad (10.16)$$

де \bar{y}_j - умовна середня значень Y j -ї групи;

N_j – об'єм j -ї групи.

Міжгрупова дисперсія – це дисперсія групових середніх відносно загальної середньої.

Внутрішньогрупова дисперсія являє собою середнє арифметичне групових дисперсій:

$$D_{\text{вн.гр}} = \frac{\sum N_j S_y}{n} \quad (10.17)$$

Можна показати, що дисперсія ознаки є сумою внутрігрупової й міжгрупової дисперсій:

$$D_y = D_{\text{вн.гр}} + D_{\text{межгр}} \quad (10.18)$$

Якщо Y функціонально залежить від X , то кожному певному значенню X відповідає єдине значення Y , і дисперсія відносно умовного середнього $S_y = 0$, відповідно дорівнює нулю середнє арифметичне дисперсій щодо умовного се-

реднього. Таким чином, при функціональному зв'язку між X і Y внутрішньо-групова дисперсія дорівнює нулю, а загальна дисперсія Y дорівнює міжгруповій дисперсії, отже, якщо зв'язок функціональний, то кореляційне відношення дорівнює одиниці:

$$\eta_{yx} = \frac{\sigma_{y_x}^2}{\sigma_y^2} = 1. \quad (10.19)$$

Якщо кореляційне відношення дорівнює нулю, $\eta=0$, то між X і Y зв'язок відсутній. Це витікає з того, що в цьому випадку міжгрупова дисперсія дорівнює нулю, а значить відсутній розкид умовних середніх відносно загальної середньої. Тобто, умовні середні при всіх значеннях X однакові, а значить Y не залежить від X.

Кореляційне відношення має наступні властивості:

- його значення лежить у межах від 0 до 1:

$$0 \leq \eta \leq 1;$$

- значення кореляційного відношення більше або дорівнює вибіркового коефіцієнту кореляції:

$$\eta \geq |r_b|;$$

4. якщо кореляційне відношення дорівнює вибіркового коефіцієнту кореляції, $\eta = |r_b|$, то між X і Y є лінійна кореляційна залежність.

Елементи регресійного аналізу

МНК дозволяє одержати точкові оцінки коефіцієнтів прийнятої залежності $Y = \varphi(X)$. Але тому, що коефіцієнти рівняння регресії – величини випадкові, вимагають перевірки й сама залежність і її коефіцієнти.

4. Перевірка адекватності рівняння регресії експериментальним даним виконується за критерієм Фішера

$$F = \frac{D_{ya}}{D_{yo}} \quad (10.20)$$

де D_{ya} – дисперсія адекватності. Вона визначається за формулою:

$$D_{ya} = \frac{1}{n - s} \sum_{i=1}^n (y_{ip} - m_{yi})^2,$$

де n – число дослідів;

s – кількість шуканих параметрів апроксимуючої залежності;

y_{ip} – розрахункове значення функції в i -й точці при апроксимації залежністю $Y = \varphi(X)$;

m_{yi} – середнє значення y в i -м досліді;

D_{yo} – дисперсія дослідів. Вона визначається на підставі даних паралельних дослідів:

$$D_{yo} = \frac{1}{m * n} \sum_{i=1}^n D_{yi},$$

де m – число паралельних дослідів в i -й точці;

n – число дослідів;

$m*n$ – загальне число вимірів;

D_{yi} – дисперсія i -го дослідів, обумовлена за формулою

$$D_{yi} = \frac{\sum_{j=1}^m (y_{ij} - m_{yi})^2}{m - 1}$$

де m_{yi} – середнє значення Y у i -м досліді.

Отримане значення F порівнюють із табличним F_T . Якщо $F < F_m$, то гіпотеза про адекватність *не відкидається*.

Тема 11. Перевірка статистичних гіпотез

Статистичні гіпотези

Будь-яка інформація, отримана в результаті обробки статистичних даних, носить імовірнісний характер. Зокрема, оцінка генеральної середньої є величиною випадковою, розподіленою нормально з параметрами \bar{x} і $\frac{\sigma}{\sqrt{n}}$. Оцінка генеральної дисперсії також випадкова. Тому будь-який висновок, заснований на статистичних даних, є науковим припущенням і називається **статистичною гіпотезою**. Статистичні гіпотези підлягають перевірці, ціль якої – визначити, чи не суперечить висунута гіпотеза вихідному статистичному матеріалу (вибірці).

Основну гіпотезу, сформульовану в результаті обробки статистичного матеріалу, називають **нульовою гіпотезою** й позначають H_0 . На протипагу нульовій гіпотезі призначають одну або декілька альтернативних (конкуруючих) гіпотез. Їх позначають H_1, H_2, \dots і т.д.

Наприклад, якщо перевіряється гіпотеза про рівність параметра a деякому заданому значенню a_0 , то як альтернативні гіпотези можна розглянути гіпотези, що a більше або менше a_0 :

$$H_0: a = a_0;$$

$$H_1: a > a_0;$$

$$H_2: a < a_0;$$

$$H_3: a \neq a_0.$$

Вибір альтернативної гіпотези обумовлюється формулюванням задачі.

В якості критеріїв для перевірки статистичних гіпотез використовують випадкові величини (статистики), особливість яких полягає в тому, що кожна з них має свій закон розподілу, що не залежить від закону розподілу генеральної сукупності й вибірки, а залежить від умов обробки вибірових даних. Значення цих випадкових величин, позначимо їх Z , з відповідними їм ймовірностями приводяться в довідкових таблицях.

На підставі вибірових даних визначають значення критерію Z і порівнюють його з табличним значенням, що відповідає умовам обробки даних. Перевірка статистичної гіпотези заснована на принципі, відповідно до якого ма-

лоймовірні події вважаються неможливими, а події, що мають більшу ймовірність, - достовірними. Якщо імовірність розрахункового значення критерію досить велика, тобто факт цілком імовірний, то говорять, що гіпотеза не суперечить даним спостереження. Якщо ж ця ймовірність мала, тобто подія практично неможлива, то говорять, що нульова гіпотеза суперечить даним спостереження, і її відхиляють.

Питання про те, яку ймовірність варто вважати досить великою або малою, вирішується не з математичних міркувань, а залежить від наслідків того, що прийнята гіпотеза виявиться невірною. Мала ймовірність, при якій значення критерію вважається практично неможливим, позначається α і називається **рівнем значущості**. У практичних задачах звичайно призначають рівень значущості $\alpha = 0,05-0,15$. Область значень критерію Z , що відповідає рівню значущості α , називають **критичною областю**. Область значень критерію Z , що відповідають імовірності $1-\alpha$, називають **областю прийняття гіпотези**. Значення критерію, що відокремлює область прийняття гіпотези від критичної області називається **критичною точкою z_k** (рис. 11.1).

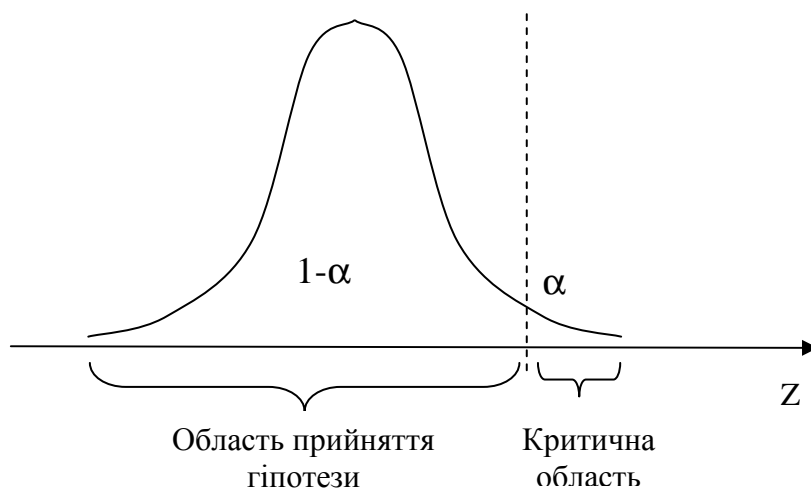


Рис. 11.1 – Розташування значень критерію Z .

Залежно від того, як сформульовані конкуруючі гіпотези, критична область може бути однією (лівосторонньою або правобічною) і двосторонньою. Відповідно критерій може мати одну або дві критичні точки (рис. 11.2).



Рис. 11.2- Критичні області

Таким чином, перевірка гіпотези заснована на факті, що критерій прийняв значення з імовірністю більшою або меншою α . Помітимо, що імовірність даного факту не дорівнює одиниці, а виходить, він не є достовірною подією. Таким чином, результат перевірки гіпотези може виявитися помилковим. Прийнято розрізняти помилки двох видів:

- 1) помилка 1-го роду – відкинута нульова гіпотеза, в той час як вона була правильною;
- 2) помилка 2-го роду – прийнята нульова гіпотеза, в той час як вона була невірною.

Якщо перевірка гіпотези показала, що вона не погодиться з вибірковими даними й повинна бути відкинута, а задача все-таки вимагає рішення, то для цього переглядають рішення задачі, використовують іншу вибірку з генеральної сукупності або збільшують обсяг вибірки. Тобто, проблема все-таки може бути вирішена.

Гірше вирішується питання, якщо зроблено помилку другого роду, тобто, прийнята невірна гіпотеза. Імовірність помилки 2-го роду позначається β . Ця ймовірність повинна бути як можна меншою. Тоді ймовірність того, що помилка 2-го роду не буде зроблена, визначиться як $1-\beta$. Величина ймовірності β залежить від якості використовуваного для перевірки гіпотези критерію. Імовірність $1-\beta$ називається **потужністю критерію**, чим вона більше, тим краще використовуваний критерій, вище надійність перевірки.

Ми розглянемо чотири критерії (нормальний розподіл, t-критерій Стьюдента, F-критерій Фішера й χ^2 -критерій Пірсона), які використовуються найбільше часто. Який з названих критеріїв варто використовувати, залежить від характеру розв'язуваної задачі, тобто від формулювання нульової гіпотези H_0 . Розглянемо ряд типових задач.

Порівняння вибіркової середньої й генеральної середньої нормальної сукупності

Нехай з нормальної генеральної сукупності витягнута вибірка об'ємом n і визначена вибіркова середня \bar{x} . Передбачається, що генеральна середня дорівнює a і генеральна дисперсія дорівнює σ^2 . Варто перевірити, чи значима розбіжність між вибірковою й генеральною середніми, або вона обумовлена випадковими причинами, тобто незначима. Запишемо нульову гіпотезу, врахуємо при цьому, що математичне сподівання вибіркової середньої дорівнює генеральній середній:

$$H_0: M[\bar{x}] = a.$$

Сформулюємо альтернативну гіпотезу:

$$H_1: M[\bar{x}] \neq a.$$

При такому формулюванні альтернативної гіпотези необхідно побудувати двосторонню критичну область, імовірність влучення в яку дорівнює рівню значимості α . Як міра розбіжності між вибірковою і генеральною середніми використаємо випадкову величину Z :

$$Z = \frac{\bar{x} - a}{\sigma_z}, \quad (11.1)$$

яка є нормованою нормальною випадковою величиною з параметрами $M[Z]=0$ і $\sigma_z = 1$.

Найбільша потужність критерію досягається, якщо імовірність влучення критерію Z у кожний із двох інтервалів критичної області дорівнює $\alpha/2$:

$$P\{|Z| > z_{кр}\} = \alpha/2. \quad (11.2)$$

Оскільки розподіл критерію Z симетричний відносно нуля, критичні точки також розташовані симетрично відносно нуля, тобто область прийняття нульової гіпотези $(-z_{кр}, z_{кр})$.

Для визначення критичних точок можна скористуватися функцією Лапласа $\Phi(x)$, що являє собою імовірність влучення нормованої випадкової величини в інтервал $(0, x)$, тобто:

$$P\{0 < X < x\} = \Phi(x).$$

Оскільки розподіл Z симетричний відносно нуля, по теоремі додавання ймовірностей маємо:

$$P\{0 < Z < z_{кр}\} + P\{z_{кр} < Z < \infty\} = 1/2,$$

або, виразивши ймовірність через функцію Лапласа, одержимо:

$$\Phi(z_{кр}) + \alpha/2 = 1/2,$$

звідки

$$\Phi(z_{кр}) = \frac{1 - \alpha}{2}. \quad (11.3)$$

Таким чином, визначивши значення функції Лапласа, можна по таблиці знайти значення її аргументу, тобто критичну точку $z_{кр}$. Тоді двостороння критична область визначається двома нерівностями:

$$Z < -z_{кр};$$

$$Z > z_{кр},$$

а область прийняття гіпотези:

$$|Z| < z_{кр}. \quad (11.4)$$

У випадку, коли генеральна дисперсія невідома, як критерій використовують t -розподіл (розподіл Стюдента) з $(n-1)$ ступенями свободи. Спостережуване значення критерію при цьому обчислюють за формулою

$$T = \frac{\bar{x} - a}{s / \sqrt{n}}, \quad (11.5)$$

де s – вибіркове середнє квадратичне відхилення.

У практичних задачах часто виникає ситуація, коли відома величина припустимої помилки δ при визначенні вибіркової середньої. Виникає задача ви-

значення об'єму вибірки, що забезпечує задану припустиму величину помилки $\delta = \bar{x} - \alpha$. Об'єм вибірки n можна визначити, скориставшись формулою (11.5). Нехай у результаті перевірки нульової гіпотези визначена двостороння критична область, що відповідає рівню значимості α , тоді

$$n = \frac{t_{\text{об'єм}}^2(\alpha, k) * s^2}{(\bar{x} - a)^2} = \frac{t_{\text{об'єм}}^2(\alpha, k) * s^2}{\delta^2}$$

де k – число ступенів свободи.

Порівняння двох дисперсій нормальних генеральних сукупностей

Нехай є дві незалежні вибірки й варто визначити, чи взяті вони з нормальних генеральних сукупностей X і Y з однаковою дисперсією. Запишемо нульову гіпотезу:

$$H_0: D[X] = D[Y], \quad (11.6)$$

і сформулюємо альтернативну гіпотезу:

$$H_1: D[X] \neq D[Y]. \quad (11.7)$$

Як критерій для перевірки нульової гіпотези про рівність дисперсій нормальних генеральних сукупностей приймають випадкову величину, що являє собою відношення більшої дисперсії до меншої:

$$F = \frac{s_1^2}{s_2^2}$$

де $s_1 > s_2$.

Нульова гіпотеза припускає, що дві вибірки незалежні й узяті з нормальних генеральних сукупностей з однаковими дисперсіями, у цьому випадку $F=1$. Однак, навіть якщо гіпотеза вірна, то малоімовірно, що s_1 прийме точно таке ж значення, як s_2 через вплив випадковості. Таким чином, завдання полягає в тому, щоб перевірити, чи буде випадкова величина F досить близька до одиниці.

Випадкова величина F має розподіл Фішера, що залежить тільки від ступенів свободи $k_1=n_1-1$ і $k_2=n_2-1$, де n_1 і n_2 об'єм вибірки з більшою й з меншою вибірковими дисперсіями відповідно, і не залежить від інших параметрів.

Для перевірки нульової гіпотези (10.6) при конкуруючій гіпотезі (10.7) будують двосторонню критичну область, що відповідає рівню значимості α . Двостороння критична область визначається двома нерівностями:

$$F < F_{\text{кр}1};$$

$$F > F_{\text{кр}2},$$

а область прийняття гіпотези:

$$F_{\text{кр}1} < F < F_{\text{кр}2}$$

причому, імовірність влучення критерію в кожний із двох інтервалів критичної області дорівнює $\alpha/2$.

Критерії згоди

Якщо по вибірці спостережень визначався закон розподілу генеральної сукупності, то виникає необхідність оцінити розбіжність між емпіричним і теоретичним розподілами. Для цього використовують *критерії згоди*, які дозволяють судити, якою є розбіжність між **емпіричним** і **теоретичним** розподілами – випадковою або значимою.

Якщо розбіжність виявиться випадковою, то вважають, що дані спостережень (вибірки) не суперечать висунутій гіпотезі про закон розподілу генеральної сукупності й, отже, гіпотезу приймають; якщо ж розбіжність виявиться значимою, то дані спостережень суперечать гіпотезі, і її відхиляють. Як міру розбіжності використовують деяку величину U .

Є декілька критеріїв згоди: критерій χ^2 (Пірсона), критерій Колмогорова, критерій Романовського і т. інше.

Пірсон запропонував як міру розбіжності використовувати суму квадратів відхилень $(p_i^* - p_i)$, узятих з деякими вагами c_i , тоді

$$U = \sum_{i=1}^k c_i (p_i^* - p_i)^2, \quad (11.8)$$

де p_i^* – частота появи ознаки X на i -м інтервалі;

p_i – теоретична ймовірність тої ж події;

k – число інтервалів;

c_i – ваговий коефіцієнт, який враховує, що відхилення $(p_i^* - p_i)$, які належать до різних груп ряду, не можна вважати рівноправними за значимістю, тому що те саме за абсолютною величиною відхилення може бути малим для великого p_i і істотним для малого p_i .

Пірсон показав, що якщо прийняти

$$c_i = \frac{n}{p_i}, \quad (11.9)$$

то величина U при збільшенні n наближається до величини χ^2 , розподіл якої залежить тільки від числа ступенів свободи

$$r = k - s, \quad (11.10)$$

де k – число інтервалів; s – число зв'язків (число незалежних умов).

Підставивши (11.9) в (11.8) і врахувавши, що $p_i^* = \frac{m_i}{n}$, дістанемо:

$$\chi_P^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}. \quad (11.11)$$

Чим більше погодяться емпіричний і теоретичний розподіли, тим менше розрізняються емпіричні й теоретичні частоти й тем менше значення χ_P^2 . Звідси витікає, що χ_P^2 характеризує близькість емпіричного й теоретичного розподілів.

Існують довідкові таблиці, в яких указана ймовірність того, що в результаті впливу випадкових факторів величина χ^2 прийме значення не менше обчисленого за даними вибірки.

Для визначеності приймають рівень значимості α (звичайно $\alpha = 0,1-0,15$). Якщо ймовірність, знайдена за таблицями, опиниться менше α , то це означає, що в результаті впливу випадкових причин наступила подія, яка практично неможлива. Таким чином, той факт, що χ^2 прийняла значення χ_P^2 не можна пояснити випадковими причинами. Це можна пояснити тим, що генеральна сукупність не розподілена за передбачуваним законом розподілу й, виходить, висунута гіпотеза про закон розподілу генеральної сукупності повинна бути відкинута. Якщо ймовірність, знайдена за таблицями, перевищує α , то гіпотеза про

закон розподілу генеральної сукупності погодиться з даними спостережень і тому може бути прийнята.

Елементи дисперсійного аналізу

Дисперсійний аналіз – метод математичної статистики, який застосовують для аналізу результатів спостережень, які залежать від різних одночасно діючих факторів. Задачі дисперсійного аналізу – вибір найбільш важливих факторів, оцінка їхнього впливу і т.інше.

Метод дисперсійного аналізу розробив англійський статистик Р.Фішер. В основі методу лежить порівняння дисперсій. На практиці дисперсійний аналіз застосовують у задачах, де потрібно оцінити вплив деякого фактору F на кількісну ознаку X. Суть дисперсійного аналізу зводиться до порівняння дисперсії, обумовленої впливом фактору F (факторної дисперсії) з дисперсією, обумовленою випадковими причинами (залишковою дисперсією). Очевидно, коли вплив фактору F є значимим, то й відмінність факторної дисперсії від залишкової дисперсії повинна бути значимою. І навпаки, якщо вплив фактору незначимий, то факторна й залишкова дисперсія відрізняються незначимо.

Нехай значення ознаки X отримані в результаті спостереження p різних груп досліду із числом спостережень в j-й групі, рівним q. Середнє значення ознаки X у кожній j-й групі (групова середня) визначиться за формулою

$$\bar{x}_{zpj} = \frac{\sum_{i=1}^q x_{ij}}{q}$$

Результати спостережень зведені в таблицю:

Номер досліду	Номер групи					
	1	2	...	j	...	p
1	x ₁₁	x ₁₂	...	x _{1j}	...	x _{1p}
2	x ₂₁	x ₂₂	...	x _{2j}	...	x _{2p}
...

i	X _{i1}	X _{i2}	...	X _{ij}	...	X _{ip}
...	
q	X _{q1}	X _{q2}	...	X _{qj}	...	X _{qp}
групова середня	\bar{x}_{zp1}	\bar{x}_{zp2}	...	\bar{x}_{zpj}	...	\bar{x}_{zpp}

Загальне число спостережуваних значень ознаки X дорівнює pq. Загальна середня визначається за формулою

$$\bar{x} = \frac{\sum_{j=1}^p \sum_{i=1}^q x_{ij}}{pq}$$

Можна оцінити повне розсіювання ознаки X, викликане як випадковими причинами, так і впливом фактору F, визначивши суму квадратів відхилень всіх спостережуваних значень x_{ij} від загальної середньої. Вона називається **загальною** сумою квадратів відхилень і визначається за формулою

$$S_{\text{общ}} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x})^2 \quad (11.12)$$

Вважають, що фактор F впливає на різні групи значень ознаки. Розсіювання за фактором або розсіювання між групами можна оцінити, визначивши суму квадратів відхилень групових середніх від загальної середньої. Її називають **факторною** сумою квадратів відхилень і визначають за формулою

$$S_{\text{факт}} = q * \sum_{j=1}^p (\bar{x}_{zpj} - \bar{x})^2 \quad (11.13)$$

Уважають, що на значення ознаки в j-й групі фактор F впливає однаково, а їхнє розсіювання обумовлене впливом випадкових причин. Суму квадратів відхилень спостережуваних значень ознаки X від своєї групової середньої \bar{x}_{zpj} називають **залишковою** сумою квадратів відхилень. Залишкова сума квадратів відхилень характеризує розсіювання всередині групи й визначається формулою:

$$S_{\text{ост}} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x}_{zpj})^2 \quad (11.14)$$

де

Можна показати, що справедливе співвідношення

$$S_{\text{общ}} = S_{\text{факт}} + S_{\text{ост}}, \quad (11.15)$$

яке найчастіше використовують для визначення залишкової суми квадратів відхилень

$$S_{\text{ост}} = S_{\text{общ}} - S_{\text{факт}}. \quad (11.16)$$

Оскільки дисперсійний аналіз припускає порівняння дисперсій, то використовуючи загальні, факторну й залишкову суми квадратів відхилень, визначають відповідні дисперсії.

Загальна дисперсія:

$$s_{\text{общ}}^2 = \frac{S_{\text{общ}}}{pq - 1} \quad (11.17)$$

де $pq - 1 = n - 1$ – число ступенів свободи загальної дисперсії.

Факторна дисперсія:

$$s_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p - 1} \quad (11.18)$$

де $p - 1$ – число ступенів свободи факторної дисперсії; p – число груп впливу фактору F .

Залишкова дисперсія:

$$s_{\text{ост}}^2 = \frac{S_{\text{ост}}}{p(q - 1)} \quad (11.19)$$

де $p(q - 1)$ – число ступенів волі залишкової дисперсії, обумовлене як різниця між числами ступенів свободи загальної й факторної дисперсій:

$$(pq - 1) - (p - 1) = p(q - 1).$$

Припустимо, що вплив фактора F відсутній. У цьому випадку групові середні $\bar{x}_{грj}$ приймають різні значення в результаті впливу тільки випадкових причин, а виходить, розрізняються незначимо. Відповідно факторна й залишкова дисперсії є незміщеними оцінками невідомої генеральної дисперсії й також розрізняються незначимо. У такій задачі формулюють нульову гіпотезу про рівність факторної й залишкової дисперсій. Якщо зрівняти оцінки цих дисперсій за критерієм F , то критерій укаже, що гіпотезу можна прийняти.

Якщо нульова гіпотеза про рівність групових середніх (а отже факторної й залишкової дисперсій) помилкова, то зі зростанням розбіжності між груповими середніми буде збільшуватися факторна дисперсія й спостережуване значення критерію F . При $F_{\text{набл}} > F_{\text{кр}}$ нульова гіпотеза про рівність факторної й залишкової дисперсій буде відкинута.

Таким чином, для того, щоб перевірити нульову гіпотезу про рівність групових середніх нормальних сукупностей з однаковими дисперсіями, варто перевірити за критерієм F нульову гіпотезу про рівність факторної й залишкової дисперсій. Причому, якщо факторна дисперсія виявиться менше залишкової, то із цього витікає справедливність гіпотези про рівність групових середніх, і F -критерій можна не обчислювати.

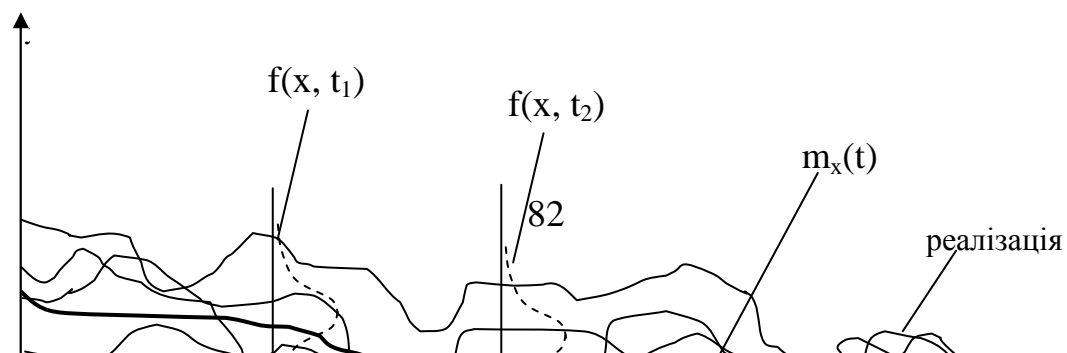
Тема 12. Елементи теорії випадкових процесів

Поняття випадкового процесу

Функція X аргументу t називається випадковою, якщо при кожному заданому значенні аргументу t величина X є випадковою. Вид, прийнятий функцією X у результаті дослідження, називається **реалізацією** функції X .

Процеси, описувані випадковими функціями, називаються **випадковими** або **стохастичними**.

На рис. 12.1 показане сімейство реалізацій випадкової функції $X(t)$.



Імовірнісні характеристики випадкового процесу є функціями часу. Якщо зафіксувати час t , то випадкова функція перетворюється в звичайну випадкову величину, що може приймати різні значення й має деякий закон розподілу зі своїми параметрами. Будемо називати цю величину перетином випадкової функції, що відповідає даному t .

Закон розподілу однієї випадкової величини є функція одного аргументу. Закон розподілу системи двох випадкових величин – функція двох аргументів і т.інше. Однак користування функціями багатьох аргументів як імовірнісні характеристики настільки незручно, що звичайно розглядають тільки їхні числові характеристики. Обмежимося розглядом найпростіших характеристик випадкових функцій, аналогічних числовим характеристикам випадкових величин. Апарат числових характеристик дозволяє порівняно просто вирішувати багато практичних задач.

На відміну від числових характеристик випадкових величин, що представляють собою певні числа, характеристики випадкових функцій являють собою не числа, а функції.

Математичним сподіванням випадкової функції $X(t)$ називається не випадкова функція, що при кожному значенні аргументу t дорівнює математичному сподіванню відповідного перетину випадкової функції:

$$m_x(t) = M[X(t)].$$

Дисперсією випадкової функції $X(t)$ називається не випадкова функція $D_x(t)$, значення якої для кожного t дорівнює дисперсії відповідного перетину випадкової функції:

$$D_x(t) = D[X(t)].$$

Таким чином, математичне сподівання – це деяка середня функція аргументу t , біля якої різним образом варіюються конкретні реалізації випадкової функції. Дисперсія випадкової функції при кожному t характеризує розкид можливих реалізацій випадкової функції щодо середнього.

Однак для опису основних особливостей випадкових функцій цих характеристик недостатньо.

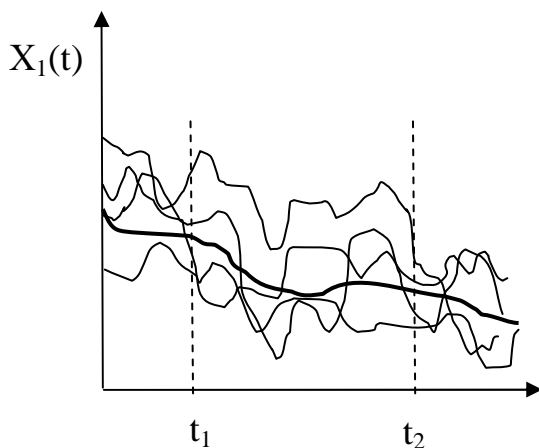


Рис. 12.2 - Сімейство реалізацій випадкової функції $X_1(t)$

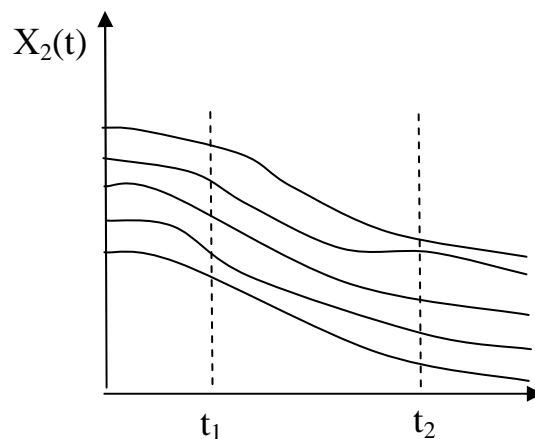


Рис. 12.3 - сімейство реалізацій випадкової функції $X_2(t)$

У випадкових функцій $X_1(t)$ і $X_2(t)$ на рис. 12.2 і 12.3 приблизно однакові математичні сподівання й дисперсії, але характер цих випадкових функцій різко відрізняється. Для $X_2(t)$ характерна яскраво виражена залежність між її значеннями при різних t (якщо при t_1 вона прийняла значення менше середнього, те ж і при t_2 - менше середнього). Випадкова функція $X_1(t)$ має різко коливальний характер з безперервними безладними коливаннями. Для такої функції характерне швидке загасання залежності між її значеннями в міру збільшення відстані по t між ними.

Внутрішня структура розглянутих процесів зовсім різна, і для її опису вводять спеціальну характеристику – кореляційну функцію. Кореляційна функція характеризує ступінь залежності між перетинами випадкової функції, що ставляться до різних t .

Кореляційною функцією випадкової функції $X(t)$ називається не випадкова функція двох аргументів $K(t, t')$, що при кожній парі значень t, t' дорівнює кореляційному моменту відповідних перетинів випадкової функції:

$$K_x(t, t') = M[\overset{0}{X}(t) * \overset{0}{X}(t')]. \quad (12.1)$$

Таким чином, якщо повернутися до малюнків 12.2 і 12.3, кореляційна функція $X_2(t)$ повільно убуває зі збільшенням проміжку (t, t') , а кореляційна функція $X_1(t)$ убуває швидко.

Якщо аргументи кореляційної функції K_x збігаються, тобто $t = t'$, то

$$K_x(t) = M[\overset{0}{X}(t) * \overset{0}{X}(t)] = M[\overset{0}{X}^2(t)] = D_x(t), \quad (12.2)$$

тобто при $t = t'$ кореляційна функція звертається в дисперсію випадкової функції $X(t)$.

Таким чином, необхідність у визначенні дисперсії випадкової функції відпадає. Як основні характеристики випадкової функції досить розглядати її математичне сподівання й кореляційну функцію.

Стаціонарний випадковий процес

На практиці часто зустрічаються процеси, що протікають у часі приблизно однорідно й мають вигляд безперервних випадкових коливань навколо деякого середнього значення. Причому, ні середня амплітуда, ні характер цих коливань не виявляють істотних змін із часом. Такі випадкові процеси називаються **стаціонарними**. Нестаціонарний процес характерний тим, що він має тенденцію розвитку в часі.

Випадковий процес $x(t)$ називається стаціонарним, якщо його імовірнісні характеристики не залежать від початку відліку часу, тобто закон розподілу перетинів той самий:

$$f(x, t_1) = f(x, t_2) = f(x).$$

Не залежать від початку відліку часу і числові характеристики – математичне сподівання і дисперсія:

$$m(t) = \text{const}; D_x(t) = \text{const}. \quad (12.3)$$

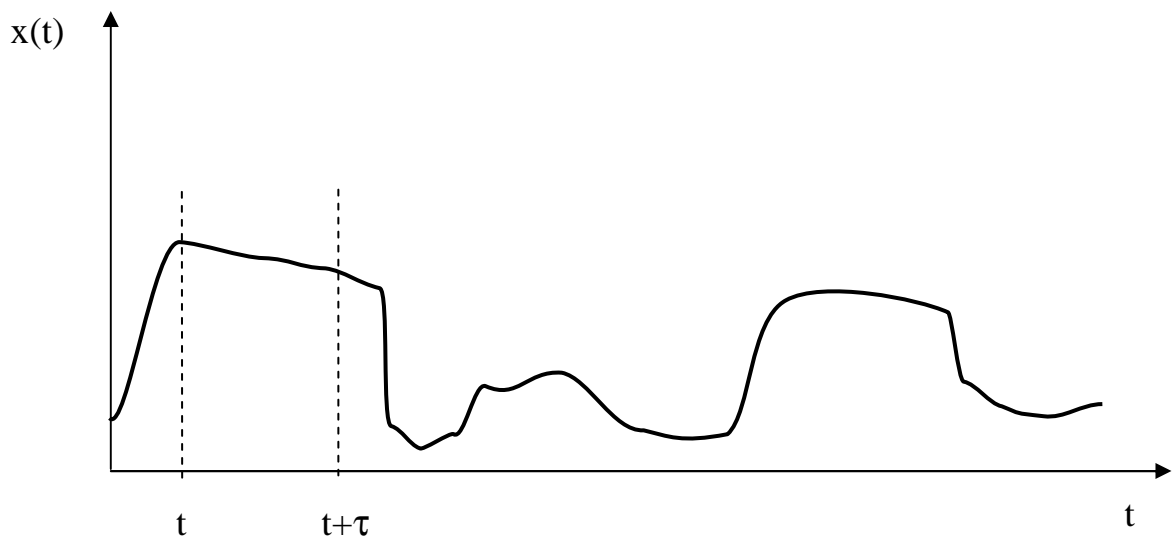


Рис. 12.4 - Стаціонарний випадковий процес

Установимо, якій умові повинна задовольняти кореляційна функція стаціонарного випадкового процесу. Покладемо $t' = t + \tau$ і розглянемо кореляційний момент двох перетинів $x(t)$, розділений інтервалом часу τ , $K(t, t + \tau)$.

Очевидно, що цей кореляційний момент не повинен залежати від того, де саме на осі взятий інтервал τ , а повинен залежати тільки від довжини цього інтервалу, тобто

$$K(t_1, t_1 + \tau) = K(\tau) = R_x(\tau). \quad (12.4)$$

Таким чином, вираження (12.4) – єдина істотна умова, якій повинен задовольняти стаціонарний випадковий процес. Ця залежність називається кореляційною функцією $R_x(\tau)$. Якщо $\tau = 0$, то кореляційна функція дорівнює дисперсії й максимальна:

$$R_X(\tau)/\tau=0 = D_X \quad (12.5)$$

Зі збільшенням τ зв'язок між перетинами стає слабкіше, тобто $R_X(0) \geq R_X(\tau)$, що показано на рис. 12.5.

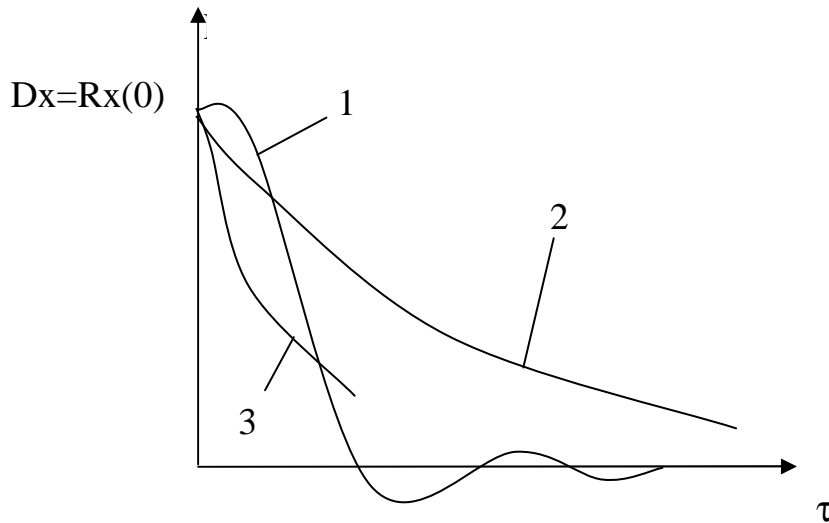


Рис. 12.5 – Кореляційна функція

Кореляційна функція характеризує внутрішній частотний склад випадкового процесу. Чим більше високочастотних складових містить випадковий процес, тим різкіше спадає зв'язок зі збільшенням τ (криві 2 і 3 на рис. 12.5). Коливання кореляційної функції (крива 1) указують на сховану періодичність процесу.

Ергодична гіпотеза

Випадковий процес називається ергодичним, якщо всі його статистичні характеристики можуть бути визначені за одною реалізацією достатньої тривалості. Стаціонарний випадковий процес звичайно має властивість ергодичності. При цьому з імовірністю, рівній одиниці, числові характеристики, визначені за множиною реалізацій,

математичне сподівання $M[x(t)] = \int_{-\infty}^{\infty} x(t) f[x(t)] dx$ (12.6)

і кореляційна функція (як другий змішаний центральний момент)

$$R_x(\tau) = M[x(t)x(t+\tau)] = \iint x(t)x(t+\tau) f(x_1, x_2) dx_1 dx_2$$
 (12.7)

дорівнюють відповідним числовим характеристикам, визначеним за часом:

математичному сподіванню $M[x(t)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt$ (12.8)

і кореляційної функції $R_x(t) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau) dt$. (12.9)

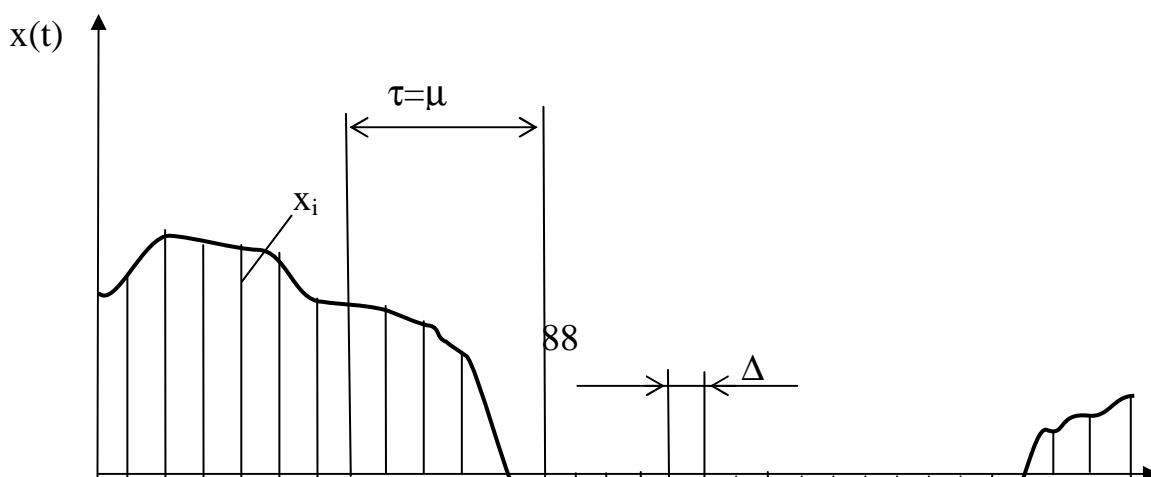
Обчислення характеристик стаціонарного ергодичного випадкового процесу по одній реалізації провадиться в наступній послідовності:

4. Для обчислення інтегралів (12.6) і (12.7) дискретизують задачу, тобто спостережуваний період часу розбивають на інтервали Δ . Інтервал дискретності вибирають на підставі теореми Котельникова:

$$\Delta \leq \frac{\pi}{\omega_{\max}}$$

де ω_{\max} – максимальна частота, що міститься в реалізації.

При цьому кількість отриманих ординат буде $N+1$, де $N=T/\Delta$.



2. Обчислюють значення математичного сподівання, замінивши інтеграл (12.8) сумою:

$$m_x = \frac{1}{N+1} \sum_{i=0}^N x_i. \quad (12.10)$$

3. Розраховують значення кореляційної функції, користуючись дискретним співвідношенням:

$$R_x(\mu) = \frac{1}{N-\mu} \sum_{i=0}^{N-\mu} x_i^0 x_{i+\mu}^0, \quad (12.11)$$

де x_i^0 - центроване значення ознаки в момент часу $t_i = i \cdot \Delta$, рівне $x_i^0 = x_i - m_x$, $i = \overline{0, N-\mu}$;

$x_{i+\mu}^0$ - центроване значення ознаки для моменту часу $t_{i+\mu}$.

Параметр μ визначають за формулою $\mu = \tau / \Delta = \overline{0, \mu_{\max}}$. Максимальне значення μ звичайно вибирають, виходячи з нижчої частоти, що міститься в реалізації:

$$\mu_{\max} = \frac{2\pi}{\omega_{\text{низ}} * \Delta} = \frac{\tau_{\max}}{\Delta}.$$

Для одержання достовірних відомостей про кореляційну функцію період спостереження T повинен набагато перевищувати τ_{\max} :

$$\tau_{\max} \leq (0,1-0,2) T \quad \text{або} \quad \mu_{\max} \leq (0,1-0,2) N.$$

4. Отримані експериментальні точки згладжують по методу найменших квадратів залежністю вигляду:

$$R_x(\tau) = Ae^{-\alpha\tau} \cos \beta\tau, \quad (12.12)$$

де визначають параметри α і β , а параметр A знаходять із (12.12) при $\mu=0$.

Елементи теорії масового обслуговування. Основні поняття.

Класифікація СМО

При дослідженні складних систем доводиться зіштовхуватися із системами, призначеними для багаторазового використання при розв'язанні однотипних задач. Виникаючи при цьому процеси отримали назву процесів обслуговування, а системи – систем масового обслуговування (СМО). Прикладами таких систем є телефонні системи, ремонтні майстерні, обчислювальні комплекси, квиткові каси, магазини, перукарні і т. інше.

Кожна СМО складається з певного числа обслуговуючих одиниць (приладів, пристроїв, пунктів, станцій), які будемо називати каналами обслуговування. Каналами можуть бути лінії зв'язку, робочі точки, обчислювальні машини, продавці і т. інше. По числу каналів СМО підрозділяють на одноканальні й многоканальні.

Заявки надходять до СМО звичайно не регулярно, а випадково, утворюють так званий випадковий потік заявок (вимог). Обслуговування заявок також триває якийсь випадковий час. Випадковий характер потоку заявок і часу обслуговування приводить до того, що СМО виявляється завантаженою нерівномірно: в якісь періоди часу накопичується дуже велика кількість заявок (вони або стають у чергу, або залишають СМО не обслугованими), в інші ж періоди СМО працює з недовантаженням або простоює.

Предметом теорії масового обслуговування є побудова математичних моделей, що зв'язують задані умови роботи СМО (число каналів, їхня продуктивність, характер потік заявок і т.п.) з показниками ефективності СМО, що описують її спроможність справлятися з потоком заявок.

Як показники ефективності СМО використовуються: середнє число заявок, що обслуговуються в одиницю часу; середнє число заявок у черзі; середній час сподівання обслуговування; імовірність відмови в обслуговуванні без

сподівання; імовірність того, що число заявок у черзі перевищить певне значення і т. інше.

СМО ділять на два основних класи: СМО з відмовами й СМО з сподіванням (із чергою). У СМО з відмовами заявка, що надійшла в момент, коли всі канали зайняті, одержує відмову, залишає СМО й надалі процесі обслуговування не бере участь (наприклад, заявка на телефонну розмову в момент, коли всі канали зайняті, одержує відмову й залишає СМО не обслугованою). У СМО з сподіванням заявка, що прийшла в момент, коли всі канали зайняті, не іде, а стає в чергу на обслуговування.

СМО з сподіванням підрозділяються на різні види залежно від того, як організована черга: з обмеженою або необмеженою довжиною черги, з обмеженим часом сподівання і т. інше.

Для класифікації СМО важливе значення має дисципліна обслуговування, яка визначає порядок вибору заявок із числа що надійшли і порядок розподілу їх між вільними каналами. За цією ознакою обслуговування заявки може бути організоване за принципом «перша прийшла – перша обслугована», «остання прийшла – перша обслугована».

Поняття марковського випадкового процесу

Процес роботи СМО являє собою випадковий процес.

Під випадковим процесом розуміється процес зміни в часі стану якої-небудь системи відповідно до імовірнісних закономірностей.

Процес називається процесом з *дискретними станами*, якщо його можливі стани S_1, S_2, \dots можна заздалегідь перелічити, а перехід системи зі стану в стан відбувається миттєво (стрибком). Процес називається процесом з *безперервним часом*, якщо моменти можливих переходів системи зі стану в стан не фіксовані заздалегідь, а випадкові.

Процес роботи СМО являє собою випадковий процес із дискретними станами й безперервним часом. Це означає, що стан СМО змінюється стрибком у випадкові моменти появи якихось подій (наприклад, приходу нової заявки, закінчення обслуговування і т. інше.).

Математичний аналіз роботи СМО істотно спрощується, якщо процес цієї роботи – марковський. Випадковий процес називається марковським або випадковим процесом без післядії, якщо для будь-якого моменту часу t_0 імовірнісні характеристики процесу в майбутньому залежать тільки від його стану в цей момент t_0 і не залежать від того, коли і як система прийшла в цей стан. Приклад марковського процесу: система S – лічильник у таксі. Стан системи в момент t характеризується числом кілометрів, що пройшов автомобіль на даний момент. Нехай у момент t_0 лічильник показує S_0 . Імовірність того, що в момент $t > t_0$ лічильник покаже те чи інше число кілометрів S_1 , залежить від S_0 , але не залежить від того, в які моменти часу змінювалися дані лічильника до моменту t_0 .

Багато процесів можна приблизно вважати марковськими, тому що в ряді випадків передісторією розглянутих процесів можна просто зневажити й застосувати для їхнього вивчення марковські моделі.

Для математичного опису марковського випадкового процесу з дискретними станами й безперервним часом, що протікає в СМО, познайомимося з одним з важливих понять – поняттям потоку подій.

Потоки подій

Під потоком подій розуміється послідовність однорідних подій, що впливають одне за іншим у якісь випадкові моменти часу (наприклад, потік викликів на телефонній станції, потік відмов ЕОМ, потік покупців і т. інше.). Потік характеризується інтенсивністю λ - частотою появи подій або середнім числом подій, що надходять у СМО за одиницю часу.

Потік подій називається **стаціонарним**, якщо його імовірнісні характеристики не залежать від часу. Зокрема, інтенсивність стаціонарного потоку є величина постійна: $\lambda(t)=\lambda$.

Потік подій називається **поток**ом без післядії, якщо для будь-яких двох непересічних ділянок часу τ_1 і τ_2 – число подій, що попадають на один з них, не залежить від числа подій, що попадають на інші.

Потік подій називається **ординарним**, якщо ймовірність падання на малий (елементарний) відрізок часу Δt двох і більше подій мала в порівнянні з імовірністю влучення однієї події.

Нагадаємо, що потік подій називається найпростішим, якщо він одночасно стаціонарний, ординарний, і не має післядії. Назва «найпростіший» пояснює тим, що СМО з найпростішими потоками має найбільш простий математичний опис. Для найпростішого потоку число m подій, що попадають на довільну ділянку часу τ , розподілене за законом Пуассона

$$P_m(\tau) = \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau},$$

для якого математичне сподівання випадкової величини дорівнює її дисперсії:

$$a = \sigma^2 = \lambda\tau.$$

Зокрема, імовірність того, що за час τ не відбудеться жодної події ($m=0$), дорівнює:

$$P_0 = e^{-\lambda\tau}.$$

Інтервал часу між довільними двома сусідніми подіями найпростішого потоку T розподілений за експонентним законом:

$$F(t) = P(T < t) = 1 - e^{-\lambda t}.$$

Щільність імовірності випадкової величини T є похідна її функції розподілу, тобто:

$$f(t) = F'(t) = \lambda e^{-\lambda t}.$$

Математичне сподівання дорівнює середньому квадратичному відхиленню випадкової величини:

$$a = \sigma = \frac{1}{\lambda}$$

і зворотне за величиною інтенсивності потоку.

Найважливіша властивість показового розподілу (присутня тільки показовому розподілу) полягає в наступному:

якщо проміжок часу, розподілений за показовим законом, уже тривав якийсь час τ , то це ніяк не впливає на закон розподілу частини, що залишилася, проміжку $(T-\tau)$: він буде таким же, як і закон розподілу всього проміжку T .

Інакше кажучи, для інтервалу часу T між двома послідовними сусідніми подіями потоку, що має показовий розподіл, будь-які відомості про те, скільки часу протікав цей інтервал, не впливають на закон розподілу частини, що залишилася. Це властивість показового закону являє собою, по суті, інше формулювання для «відсутності післядії» – основної властивості найпростішого потоку.

СМО з відмовами. Сталий режим обслуговування

Розглянемо функціонування системи з відмовами. Нехай є n -канальна система, яку можна представити як фізичну систему з кінцевою множиною станів: S_0 – всі канали вільні, S_1 – зайнятий рівно один канал, S_k – зайнято рівно k каналів, ..., S_n – зайняті всі n каналів. Схема можливих станів системи показана на рис 12.7:

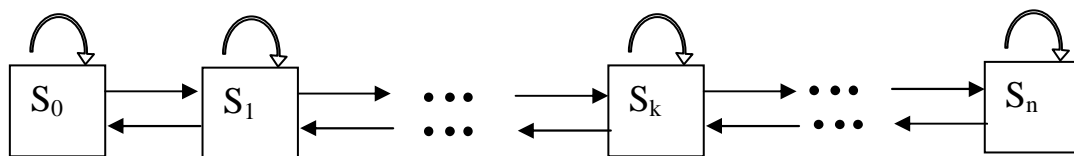


Рис.12.7 - Схема можливих станів системи з відмовами

Імовірності станів системи $p_k(t)$ для будь-якого моменту часу t можна визначити в такий спосіб. Нехай потоки заявок на обслуговування й визволення є

найпростішими, і інтервал між заявками розподілений за експонентним законом з параметром λ :

$$f(t) = \lambda e^{-\lambda t}.$$

А час перебування заявки на обслуговуванні також розподілено за експонентним законом з параметром μ :

$$g(t) = \mu e^{-\mu t}.$$

Тоді процес зміни станів системи буде марковським, що дозволяє скласти лінійні диференціальні рівняння відносно ймовірностей станів системи.

Складемо рівняння для $p_k(t)$ ($1 \leq k \leq n$). Зафіксуємо момент часу t і знайдемо ймовірність $p_k(t+\Delta t)$ того, що в момент часу $t+\Delta t$ система буде в стані S_k . Ця ймовірність обчислюється як ймовірність суми трьох подій A , B і C (за числом стрілок, що спрямовані у стан S_k на рис. 12.7): подія A = {в момент t система перебувала в стані S_k і за час Δt заявка не надійшла й жоден канал не звільнився}; подія B = {у момент t система перебувала в стані S_{k-1} , і за час Δt надійшла одна заявка}; подія C = {у момент t система перебувала в стані S_{k+1} , і за час Δt канал звільнився}. За теоремою додавання ймовірностей маємо:

$$p_k(t+\Delta t) = P(A) + P(B) + P(C). \quad (12.13)$$

Ймовірність події A , тобто того, що за час Δt не надійшла жодна заявка й жоден канал не звільнився, знайдемо відповідно до теореми множення ймовірностей:

$$[1 - F(t)] * [G(t)]^k = e^{-\lambda \Delta t} (e^{-\mu \Delta t})^k = e^{-(\lambda + \mu k) \Delta t}.$$

Нехтуючи величинами малих порядків, маємо:

$$e^{-(\lambda + \mu k) \Delta t} \approx 1 - (\lambda + \mu k) \Delta t,$$

тоді

$$P(A) = p_k(t) [1 - (\lambda + \mu k) \Delta t]. \quad (12.14)$$

Аналогічно визначимо ймовірності подій B і C :

$$P(B) = p_{k-1}(t) \lambda \Delta t, \quad (12.15)$$

$$P(C) = p_{k+1}(t) (k+1) \mu \Delta t. \quad (12.16)$$

Підставивши (12.14), (12.15) і (12.16) в (12.13), дістанемо

$$p_k(t + \Delta t) = p_k(t)[1 - (\lambda + \mu k)\Delta t] + p_{k-1}(t)\lambda\Delta t + p_{k+1}(t)(k + 1)\mu\Delta t .$$

Перенесемо $p_k(t)$ у ліву частину рівняння, розділивши на Δt і перейшовши до границі при $\Delta t \rightarrow 0$, дістанемо диференціальне рівняння для $p_k(t)$:

$$\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)p_k(t) + (k + 1)\mu p_{k+1}(t) \quad (12.17)$$

Аналогічним способом можна дістати диференціальні рівняння для ймовірностей станів системи при $k=0$ і $k=n$:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t) , \quad (12.18)$$

$$\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - n\mu p_n(t) . \quad (12.19)$$

Система рівнянь (12.17) називається рівняннями Колмогорова. Інтегрування цієї системи за початкових умов $p_0(0) = 1, p_1(0) = \dots = p_n(0) = 0$ (у початковий момент часу всі канали вільні) дозволяє знайти ймовірність кожного з $k = \overline{0, n}$ станів системи. Імовірність $p_n(t)$ – це ймовірність того, що заявка, що прийшла в момент часу t , застане всі канали зайнятими, тобто одержить відмову.

Імовірність $q(t) = 1 - p_k(t)$ називається відносною пропускною спроможністю системи. Для даного t це є відношення середнього числа обслугованих за одиницю часу заявок до середнього числа поданих.

Рівняння Колмогорова дають можливість знайти всі ймовірності станів як функції часу. Особливий інтерес представляють ймовірності системи $p_k(t)$ у граничному стаціонарному режимі, тобто при $t \rightarrow \infty$, які називаються граничними (або фінальними) ймовірностями станів.

Теорія випадкових процесів доводить, що якщо число станів системи скінченне й з кожного з них можна (за кінцеве число кроків) перейти в будь-який інший стан, то граничні ймовірності існують.

Гранична ймовірність стану S_k , має чіткий сенс: вона показує середній відносний час перебування системи в цьому стані. Наприклад, якщо гранична

ймовірність стану S_0 , тобто $p_0=0,5$, то це означає, що в середньому половину часу система перебуває в стані S_0 .

Оскільки граничні ймовірності постійні, то, замінюючи в рівняннях Колмогорова їхні похідні нульовими значеннями, дістанемо систему лінійних алгебраїчних рівнянь, що описують стаціонарний режим:

$$\begin{cases} -\lambda p_0 + \mu p_1 = 0 \\ \dots \\ \lambda p_{k-1} - (\lambda + k\mu) p_k + (k+1)\mu p_{k+1} = 0. \\ \dots \\ \lambda p_{n-1} - n\mu p_n = 0 \end{cases} \quad (12.20)$$

Розв'язавши систему (12.20) щодо ймовірностей, отримаємо для будь-якого k :

$$p_k = \frac{\lambda^k}{k! \mu^k} p_0. \quad (12.21)$$

Позначимо $\lambda/\mu=\alpha$ і назвемо її приведеною інтенсивністю потоку заявок, що є ні що інше, як число заявок, що доводиться на середній час обслуговування однієї заявки. При цьому формула (12.21) матиме вигляд:

$$p_k = \frac{\alpha^k}{k!} p_0. \quad (12.22)$$

Звідки, урахувавши, що $\sum_{k=0}^n p_k = 1$, одержимо:

$$\sum_{k=0}^n p_k = p_0 \sum_{k=0}^n \frac{\alpha^k}{k!} = 1.$$

Виразимо звідси p_0 , підставимо в формулу (12.22) та отримаємо:

$$p_k = \frac{\alpha^k}{\sum_{k=0}^n \frac{\alpha^k}{k!}}, \text{ де } k = \overline{0, n}. \quad (12.23)$$

Формула (12.23) називається формулою Ерланга. Вона дає граничний закон розподілу числа зайнятих каналів залежно від характеристик потоку заявок і продуктивності системи обслуговування.

Помітимо, що незважаючи на те, що формули Ерланга в точності справедливі тільки для найпростішого потоку заявок, ними з відомим наближенням можна користуватися й для потоків з обмеженою післядією, а також для систем

з очікуванням, коли строк очікування заявки в черзі малий у порівнянні із середнім часом обслуговування.

СМО з очікуванням. Метод статистичних випробувань

Якщо заявка, що застала всі канали зайнятими, стає в чергу й чекає, поки звільниться який-небудь канал, причому час сподівання не обмежений, то така система називається чистою системою з очікуванням. Якщо час очікування обмежений, то система називається системою змішаного типу. Обмеження можуть накладатися на час очікування заявки в черзі, на число заявок у черзі. При цьому заявки, що очікують, можуть викликатися на обслуговування за чергою, у випадковому або пільговому порядку. Кожний тип системи має свої особливості й свою методику розрахунку. Наприклад, змішані системи деяких типів можуть бути описані диференціальними рівняннями, аналогічними рівнянням Эрланга. При цьому час очікування заявки в черзі, як і час її обслуговування, вважається випадковим, розподіленим за експонентним законом. Це дозволяє вважати процеси, що протікають у системі, марковськими.

У тих випадках, коли випадкові процеси не можна віднести до класу марковських, аналітичне вираження для розрахунку ймовірностей станів системи одержати не вдається. У цьому випадку користуються універсальним методом статистичних випробувань або, як його часто називають, методом Монте-Карло.

У першому наближенні ідею методу можна описати так: процес функціонування складної системи імітується на ЕОМ з усіма супровідними його випадками. Таким чином, будується одна реалізація випадкового процесу з випадковим ходом і результатом. Сама по собі така реалізація не дає підстав до вибору рішення, але, одержавши множину таких реалізацій, можна їх обробити як звичайний статистичний матеріал, визначити середні характеристики по множині й одержати подання про те, які умови задачі й елементи рішення впливають на функціонування системи. Таким чином, при використанні методу Монте-Карло випадковість використовується як апарат дослідження.

Для систем, у яких бере участь велика кількість різнорідних елементів, а

випадкові фактори складним способом взаємодіють між собою, процеси неможливо описати аналітично як за допомогою детермінованих, так і ймовірнісних методів, а імітаційне моделювання, як правило, виявляється простіше аналітичного, а нерідко і єдино можливим методом моделювання.

СПИСОК ЛІТЕРАТУРИ

1. Гмурман В. Э. Теория вероятностей и математическая статистика. – М.: Высш. школа, 1977. – 498 с.
2. Гмурман В.Э. Руководство к решению задач по теории вероятностей и математической статистике. – М.: Высш. школа, 1975. – 330с.
3. Вентцель Е. С. Теория вероятностей. – М.: Высшая школа, 1999.
4. Гнеденко Б.В. Курс теории вероятностей.- Физматгиз, 1961.
5. Вентцель Е. С., Овчаров Л.А. Сборник задач по теории вероятностей. – М.: Наука, 1969.
6. Сборник задач по теории вероятностей, математической статистике и теории случайных функций./ Под ред. А.А.Свешникова.- Наука, 1970.

ЗМІСТ

ВСТУП	3
ЗМ 1. ІМОВІРНІСТЬ ВИПАДКОВОЇ ПОДІЇ.....	7
Тема 1. Основні поняття теорії ймовірностей.....	7
Класичний і статистичний методи визначення ймовірності випадкової події	8
Тема 2. Операції над подіями. Теореми теорії ймовірностей Основні формули теорії ймовірностей.	10
Теорема додавання.....	11
Теорема множення	13
Формула повної ймовірності.....	14
Формула Бейеса (теорема гіпотез)	14
Формула Бернуллі	15
Локальна теорема Лапласа.	16
Інтегральна теорема Лапласа.	16
Формула Пуассона.	17
ЗМ 2. ВИПАДКОВІ ВЕЛИЧИНИ І ЇХНІ ЗАКОНИ РОЗПОДІЛУ	18
Тема 3. Поняття випадкової величини і закону розподілу. Універсальні закони розподілу	18
Функція розподілу.....	19
Щільність розподілу	20
Тема 4. Числові характеристики випадкової величини.	22
Властивості числових характеристик	25
Тема 5. Найбільш важливі для практики закони розподілу випадкових величин.....	26
Біноміальний закон розподілу	26
Закон розподілу Пуассона.....	28
Експонентний закон розподілу	29
Нормальний закон розподілу ймовірностей.....	31

Поняття про центральну граничну теорему	34
Закон рівномірної щільності.	35
Тема 6. Система випадкових величин. Закони розподілу й числові характеристики системи.	37
Багатомірна випадкова величина	37
Функція розподілу системи двох випадкових величин	37
Щільність розподілу системи двох випадкових величин	38
Числові характеристики системи випадкових величин	38
Функції випадкових величин	40
Тема 7. Закон великих чисел.	43
Принцип практичної впевненості. Формулювання закону великих чисел	43
ЗМ 3. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ.	49
Тема 8. Основні поняття. Суть вибіркового методу.	49
Визначення закону розподілу спостережуваної ознаки за статистичними даними	54
Тема 9. Числові характеристики варіаційного ряду	56
Властивості вибірових числових характеристик	58
Довірчий інтервал і довірча імовірність	60
Тема 10. Елементи теорії кореляції.	62
Метод найменших квадратів.	64
Вибірковий коефіцієнт кореляції	66
Вибіркове кореляційне відношення	68
Елементи регресійного аналізу	69
Тема 11. Перевірка статистичних гіпотез.	71
Статистичні гіпотези.	71
Порівняння вибіркової середньої й генеральної середньої нормальної сукупності.	74
Порівняння двох дисперсій нормальних генеральних сукупностей .	76
Критерії згоди	77

Елементи дисперсійного аналізу	79
Тема 12. Елементи теорії випадкових процесів	82
Поняття випадкового процесу	82
Стационарний випадковий процес	85
Ергодична гіпотеза.....	88
Елементи теорії масового обслуговування. Основні поняття.	
Класифікація СМО.....	90
Поняття марковського випадкового процесу.....	92
Потоки подій.....	93
СМО з відмовами. Сталий режим обслуговування	94
СМО з очікуванням. Метод статистичних випробувань	98
СПИСОК ЛІТЕРАТУРИ.....	100

Навчальне видання

Ачкасов Анатолій Єгорович,
Воронков Олексій Олександрович,
Воронкова Тетяна Борисівна

Конспект лекцій

з курсу

«ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА»

(для студентів 2 курсу ФПО і ЗН напряму підготовки
(0921) 6.060101 «Будівництво» та слухачів другої вищої освіти
за спеціальністю «Теплогазопостачання і вентиляція»)

Відповідальний за випуск *А. Є. Ачкасов*

Редактор *Д. Ф. Курильченко*

План 2007, поз. 167 Л

Підп. до друку 20.11.07
Друк на ризографі.
Зам. №

Формат 60x84 1/16
Ум. друк. арк. 6,1
Тираж 100 пр.

Видавець і виготовлювач:
Харківська національна академія міського господарства
вул. Революції, 12, Харків, 61002
Електронна адреса: rectorat@ksame.kharkov.ua
Свідоцтво суб'єкта видавничої справи:
ДК № 731 від 19.12.2001