МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Донецький національний університет економіки і торгівлі
імені Михайла Туган-Барановського

Кафедра загальноінженерних дисциплін та обладнання

В.М. Серебреников, Т.В. Квітка

МАТЕМАТИКА ДЛЯ ЕКОНОМІСТІВ: ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Навчально-методичний посібник для вивчення дисципліни

розділ «Математична статистика»

Кривий Ріг

2019р.

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Донецький національний університет економіки і торгівлі
імені Михайла Туган-Барановського

Кафедра загальноінженерних дисциплін та обладнання

В.М. Серебреников, Т.В. Квітка

МАТЕМАТИКА ДЛЯ ЕКОНОМІСТІВ: ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Навчально-методичний посібник для вивчення дисципліни

розділ «Математична статистика»

Затверджено на засіданні
кафедри загальноінженерних
дисциплін та обладнання
Протокол № 5
від “28” листопада 2019 р.

Схвалено навчально-методичною радою
ДонНУЕТ ім. М. Туган-Барановського
Протокол № 4
від “23” листопада 2019 р.

Кривий Ріг

2019р

УДК 591.21 (075.8)

Рекомендовано до видання науково-методичною радою Донецького національного університету економіки і торгівлі ім. Михайла Туган-Барановського (протокол № 4 від 23 листопада 2019р.)

Рецензенти:

Слободянюк Н.О., д.е.н., доцент, завідувач кафедри фінансів та банківської справи Донецького національного університету економіки і торгівлі ім. Михайла Туган-Барановського.

Омельченко О.В., к.т.н., доцент, завідувач кафедри загальноінженерних дисциплін та обладнання Донецького національного університету економіки і торгівлі ім. М. Туган-Барановського.

Серебреников В.М., Квітка Т.В.

Математика для економістів:теорія ймовірностей і математична статистика. Розділ «Математична статистика» [Текст]: навч.-метод. посібник та завдання для виконання самостійної роботи студентів денної та заочної форм навчання / В.М.Серебреников, Т.В. Квітка; М-во освіти і науки України, Донец. нац. ун-т економіки і торгівлі ім. М.Туган-Барановського, каф. загальноінженерних дисциплін та обладнання. – Кривий Ріг: ДонНУЕТ, 2019. – 76 с.

Запропонований навчально-методичний посібник є продовженням розділів «Випадкові події» і «Випадкові величини» навчально-методичних посібників, які вийшли раніше. Навчально-методичний посібник містить короткі теоретичні відомості з кожної теми, питання, що актуалізують знання, отримані під час лекцій, приклади розв’язування задач, задачі для самостійного розв’язування, які можна використати як завдання для контрольних робіт, 30 варіантів індивідуальних та тестових завдань. Рекомендується для використання у процесі навчання математики для економістів при вивченні змістового модуля «Теорія ймовірностей і математична статистика».

Наведено список рекомендованої літератури.

УДК 591.21 (075.8)

З М І С Т

ВСТУП

Навчально-методичний посібник пропонується для роботи на практичних заняттях, а також для самостійної підготовки студентів до практичних занять з навчальної дисципліни «Математика для економістів: теорія ймовірностей і математична статистика».

Передбачається, що самостійна робота з використанням розроблених начально-методичного посібника допоможе якісно підготуватись до практичних занять та модульних контрольних робіт, розвиватиме вміння та навички самостійної діяльності.

Наразі надається для самостійного опрацювання матеріал до практичних занять другого модуля («Математична статистика») відповідно робочої програми навчальної дисципліни «Математика для економістів: теорія ймовірностей і математична статистика».

В результаті вивчення дисципліни студент повинен

знати:

– поняття генеральної і вибіркової сукупності,
– способи відбору,
– статистичний розгляд вибірки,
– емпірічна функція розподілу,
– статистична сукупність, гістограма,
– числові характеристики статистичногорозподілу,
– довірчий інтервал, довірча ймовірність,
– розподіл Стьюдента,
– визначення наближених значень числових характеристик системи двох випадкових величин,
– метод найбільшої правдоподібності для знаходження оцінок параметрів розподілів,
– метод наименьших квадратов,
– статистична перевірка гіпотез,
– поняття про критерії згоди,

вміти:

– формувати вибіркові сукупності з подольшою побудовою емпирічних функцій розподілу та гістограм,
– обчислювати статистичні характеристики випадкових величин з використанням відповідних означень і теорем,

мати навички:

– самостійної роботи при розширенні своїх математичних знань та освоєння довідкових систем.

Порядок поточного оцінювання знань студентів з дисципліни

Успішне виконання студентом завдань поточного контролю є обов'язковою умовою участі його в складанні екзамену. Об'єктом поточного контролю знань студента є :

- контроль систематичності та активності роботи протягом семестру протягом семестру над вивченням програмного матеріалу дисципліни,
- виконання завдань для самостійного опрацювання,
- контроль за виконанням типових розрахунків.

Результати виконаної роботи подаються студентами в окремому зошиті.

Завдання з типових розрахунків виконуються студентом в позааудиторний час протягом вивчення дисципліни і здаються в указаний викладачем термін. До захисту типового розрахунку студент допускається при умові правильного виконання всіх завдань.

«Навчально-методичний посібник» буде корисними студентам усіх форм навчання при підготовці до аудиторних занять, у процесі самоосвіти та при підготовці до контрольних робіт та іспиту з навчальної дисципліни.

ЧАСТИНА 4
НАВЧАЛЬНО-МЕТОДИЧНІ МАТЕРІАЛИ З ПІДГОТОВКИ ДО ПРАКТИЧНИХ ЗАНЯТЬ ПО РОЗДІЛУ «МАТЕМАТИЧНА СТАТИСТИКА»

1. Предмет математичної статистики

Математичною статистикою називається наука, що займається розробкою методів отримання, опису таобробки дослідних даних з метою вивчення закономірностейвипадкових масових явищ.

Визначення методів обробки дослідних даних становить одну з основних прикладних задач теорій ймовірностей.

Всі завдання математичної статистики стосуються питань обробки спостережень над масовими випадковими явищами, але в залежності від характеру вимірюваної величини, цілі вимірювання при обробці результатів вимірювань ці завдання можуть приймати ту чи іншу форму. Типовими завданнями математичної статистики, які найбільш важливі для нас по своїм практичним застосуванням, є наступні.

1.Оцінка на підставі результатів вимірювань невідомої функції розподілу. Завдання ставиться так: в результаті незалежних вимірювань (випробувань) над випадковою величиною X отримані наступні її значення: x_1, x_2, ..., x_n.

Потрібно приблизно оцінити невідому функцію розподілу F(x) випадкової величини X.

2. Оцінка невідомих параметрів розподілу. Завдання ставиться так: випадкова величина X має функцію розподілу певного виду, що залежить від k параметрів, значення яких невідомі (про тип функціїрозподілу часто можна зробити досить певний висновок на підставі загальнотеоретичних міркувань).

Потрібно на підставі експериментальних данихоцінити значення цих параметрів.

3. Статистична перевірка гіпотез. Одна зосновних завдань статистичної перевірки гіпотез ставитьсятак: на підставі деяких міркувань можнавважати, що функція розподілу досліджуваноївипадкової величини X є F (x).Питається: чи сумісні спостережені значення з гіпотезою, що випадкова величина X дійсно має розподіл F (x).

Зокрема, якщо закон розподілу досліджуваної випадкової величини X не викликає сумнівів і в перевірці потребують тільки значення деяких параметрів, що характеризують розподіл, то в задачі питається: не спростовують чи досвідчені дані ту гіпотезу, що параметри закону розподілу мають припущенні значення.

В основі математичної статистики лежить ряд понять, без попереднього ознайомлення з якими неможливе вивчення сучасних методів обробки дослідних даних. Зупинимося на виявленні сутності основних понять математичної статистики.

2. ГЕНЕРАЛЬНА СУКУПНІСТЬ І ВИБІРКА

Нехай потрібно дослідити яку-небудь ознаку, властиву великій групі однотипних виробів (наприклад, кількість, вага виробів і т. д.).Сукупність значень ознаки всіх N виробівданого типу називається **генеральною сукупністю**. При цьому

передбачається, що число N у генеральній сукупностідуже велике. На практиці, однак, суцільне обстеження застосовується порівняно рідко. Наприклад, якщо сукупність містить дуже велике число виробів, то провести суцільне обстеження фізично невозможно.Тем більш, якщо обстеження виробів пов'язано з їх знищенням (наприклад, перевірка на якість продуктів харчування) або вимагає великих матеріальних витрат, то проводити суцільне обстеження практично не має сенсу. У таких випадках випадково відбирають з усієї сукупності обмежене число об'єктів (виробів) і піддають їх вивченню.

Вибіркової сукупністю, або просто **вибіркою**,називають сукупність випадково відібраних об'єктів.Таким чином, вибірковий метод полягає в тому,що з генеральної сукупності береться вибіркаобсягу п (причому $n << N$) і визначаються характеристики вибірки, які приймаються в якості наближених значень відповідних характеристик генеральної сукупності.

Чим більше n, тим більше обгрунтоване судження можна висловити на основі вибірки про властивості генеральної сукупності. Очевидно, що при $n \to N$ вибіркове розподіл наближається до генерального. Відзначимо, що вибірка дає найбільшу інформацію про генеральної сукупності тільки в тому випадку, коли результати обстежень, складові вибірку, є незалежними.

3.СТАТИСТИЧНИЙ РЯД. СТАТИСТИЧНА ФУНКЦІЯРОЗПОДІЛУ

Припустимо, що вивчається деяка випадкова величина X, закон розподілу якої невідомий. З цією метою над випадковою величиною X робbться ряд незалежних дослідів (вимірювань). Результати вимірювань представляють у вигляді таблиці, що складається з двох рядків, в першому з яких вказуються номера вимірювань і, а в другій - результати вимірювань x_i

i	1	2	3	…	n
x_i	x_1	x_2	x_3	…	x_n

Таблицю, в якій містяться номери і результати вимірювань, в математичній статистиці називають **статистичним рядом**. Статистичний ряд представляє собою первинну форму записи статистичного матеріалу і може бути оброблений різними способами. Одниміз способів такої обробки є побудова статистичної функції розподілу випадкової величини X.

Статистичної функцією розподілу випадкової величини називається закон зміни частоти події $X < x$ в даному статистичному матеріалі:

$$F^*(x) = P^*(X < x).$$

Для того, щоб знайти значення статистичної функції розподілу при даному x, треба підрахувати число дослідів, в яких випадкова величина X прийняла значення, менші, ніж x, і розділити на загальне числозроблених дослідів.

Статистична функція розподілу будь-якої випадкової величини (дискретної або неперервної) представляє завжди ступінчасту функцію, скачки якої відповідають спостережуваним значенням випадковоївеличини і за величиною рівні частотам цих значень.Статистична функція розподілу $F^*(x)$ призбільшенні п сходиться по ймовірності до справжньої функціїрозподілу $F(x)$ випадкової величини X.

Слід зауважити, що при великій кількості дослідів nпобудова статистичної функції розподілу $F^*(x)$ дуже трудомістка операція, тому часто буває зручно користуватися іншими характеристиками статистичнихрозподілів, аналогічними не функції розподілу $F(x)$, а щільності ймовірності $f(x)$.

4. СТАТИСТИЧНА СУКУПНІСТЬ. ГІСТОГРАМА

При великій кількості спостережень представлення результатів спостережень у вигляді статистичного ряду викликає труднощі, а при вирішенні багатьох завдань інедоцільним. У таких випадках проводять підрахунокрезультатів спостережень, що потрапляють в певні групи,і становлять таблицю, в якій вказуються групиі частота отримання результатів спостережень в кожній групі. Сукупність груп, на які розбиваютьсярезультати спостережень і частот отримання результатівспостережень в кожній групі, називають **статистичної сукупністю**. Статистична сукупність утворюється з статистичного ряду шляхом ділення його на групи по деяких ознаках і підрахунку чисел і частот вимірювань в кожній групі.

Зауважимо, що якщо при угрупованні маємо значення, яке в точності лежить на кордоні двох груп, то слід додати до чисел m_i однієї й іншої груп по 0.5. Що стосується числа груп, то їх кількість вибирається таким чином, щоб результати вимірювань були доступні для огляду імістили досить велику кількість відомостей.

Графічним зображенням статистичної сукупностіє **гістограмма**.Гістограмма будується наступним чином: по осі абсціссоткладиваются інтервали,що відповідають групам сукупності, і на кожному з них, якна підставі, будується прямокутник, площа якого дорівнює частоті даної групи. З способу побудови гістограми випливає, що повна площа її дорівнює одиниці.

Очевидно, що якщо точки гістограми з'єднати плавною лінією, то ця лінія в першому наближенні буде представляти графік щільності ймовірності випадкової величини X. При цьому, якщо число дослідів збільшувати і вибирати більш дрібні групи в статистичної сукупності, то гістограма буде все більше наближатися до щільності ймовірності випадкової величини.

Користуючись даними статистичної сукупності, можна наближено побудувати і статистичну функцію розподілу випадкової величини X.

5. ЧИСЛОВІ ХАРАКТЕРИСТИКИ СТАТИСТИЧНОГО РОЗПОДІЛУ

Закон розподілу випадкової величини являє собою деяку функцію, зазначення цієї функції повністю описує випадкову величину з імовірнісної точки зору. Однак при вирішенні багатьох практичних завдань немає необхідності характеризувати випадкову величину вичерпним чином, а достаточноуказать тільки окремі числові характеристики, які характеризують істотні риси распределеніяслучайной величини. Основними числовими характеристиками випадкової величини є математичне сподівання і дисперсія. Математичне сподівання характерізуетсереднє значення, біля якого групуються можливі значення випадкової величини, а дисперсія характеризує ступінь розкиданості цих значень відносно середнього.

Аналогічні числові характеристики існують і для статистичних розподілів. Аналогією математичного сподівання випадкової величини X є середнє арифметичне спостережуваних значень випадкової величини

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

де x_i - значення випадкової величини, що спостерігається в і-м досліді;

n - число дослідів.

Цю характеристику називають статистичними середнім випадкової величини.

При великому числі дослідів середнє арифметичне спостережуваних значень випадкової величини наближається(збігається по ймовірності) до її математичного сподівання і може бути прийнято приблизно рівним математичному сподіванню.

Аналогією дисперсії випадкової величини X є статистична дисперсія, яка визначається наступним чином

$$D^*[X]=\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

де x_i - значення випадкової величини, що спостерігається в і-м досліді;

n - число дослідів.

Аналогічно визначаються статистичні початковіі центральні моменти будь-яких порядків

$$M^*[X^k] = \frac{1}{n}\sum_{i=1}^{n} x_i^k,$$

$$M^*[(X-\bar{x})^k]=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^k$$

Зауважимо, що при збільшенні числа спостережень всестатистичні характеристики будуть сходитися по ймовірності до відповідних числовим характеристик випадкової величини і при достатньому n можуть бути прийняті наближено рівними їм.

6. ВЛАСТИВОСТІ ТОЧКОВИХ ОЦІНОК

Розглянемо наступну загальну задачу. Є випадкова величина X, закон розподілу якоїмістить невідомий параметр a. Потрібно на підставі даних знайти підходящу оцінку параметра a.

Позначимо через

$$X_1, X_2, ..., X_n$$

спостережувані значення випадкової величини X в результаті проведених n незалежних дослідів. Нехай величина $\tilde{a}$, обчислена на основі матеріалу, є оцінкою параметра a. Це означає, що $\tilde{a}$ є функцією величин $X_1, X_2, ..., X_n$:

$$\tilde{a} = \tilde{a}(X_1, X_2, ..., X_n).$$

Крім того, спостерігаються значення $X_1, X_2, ..., X_n$ слід розглядати як випадкові величини, кожна з яких розподілена за тим же законом, що і випадкова величина X. Тому $\tilde{a}$ є теж випадковою величиною, закон розподілу якої залежить, по-перше, від закону розподілу випадкової величини X, по-друге, від числа дослідів n. Для того.чтоби оцінка $\tilde{a}$ мала практичну цінність, вона повинна мати наступні властивості.

1. **Незміщеність оцінки**. Розрізняють оцінки зміщені і незміщені. Зміщеними називаються оцінки, математичне сподівання яких не дорівнює оцінюваному параметру:

$$M[\tilde{a}(X_1, X_2, ..., X_n) \neq a.$$

Незміщеними називають оцінки, для яких виконується умова

$$M[\tilde{a}(X_1, X_2,..., X_n)] = a.$$

Природно в якості наближеного невідомого параметра брати незміщені оцінки, для того щоб не робити систематичної помилки в бік завищення або заниження.

2. Спроможність оцінки. Оцінка $\tilde{a}(X_1, X_2, ..., X_n)$для параметра а називається спроможною, якщо вонасходиться по ймовірності до оцінюваного параметру принеобмеженому зростанні числа дослідів n , тобто

$$\lim_{n\to\infty} P[|\tilde{a}(X_1, X_2, \dots X_n) - a| < \varepsilon] = 1$$

де ε -скіль завгодно мале додатнє число.

Для задоволення цієї вимоги досить, щобдисперсія оцінки прагнула до нуля при $n\to\infty$, тобто, щоб виконувалася умова

$$\lim_{n\to\infty} D[|\tilde{a}(X_1, X_2, \dots X_n)|] = 0$$

і, крім того, щоб оцінка була незміщеною.

Отже, спроможність оцінки означає, що придосить великій кількості дослідів n з як завгодновеликою вірогідністю відхилення оцінки від істинногозначення параметра менше будь-який наперед заданоївеличини.Очевидно, такій вимозі повинна задовольнятивсяка оцінка, придатна для практичноговикористання.

3. Ефективність оцінки. Оцінки, що володіють властивістю незміщеності і спроможності, при обмеженій кількості дослідів можуть відрізнятися дисперсиями.Цілком очевидно, що чим менше дисперсія оцінки,тим менше ймовірність грубої помилки при визначеннінаближеного значення параметра. Тому необхідно,щоб дисперсія оцінки була мінімальною, тобто щобвиконувалася умова

$$D[\tilde{a}(X_1, X_2, \dots, X_n)] = D_{min}$$

Оцінка, що має таку властивість, називається **ефективною**.

При виробленні практичних методів обробки дослідних даних з метою отримання оцінок, прийнятих в якості наближених значень шуканих параметрів, необхідно керуватися сформульованими властивостями оцінок.

7. ВИЗНАЧЕННЯ НАБЛИЖЕНИХ ЗНАЧЕНЬ ВИМІРЮВАНОЇ ВЕЛИЧИНИ І ДИСПЕРСІЇ У ВИПАДКУ ПРЯМИХ РІВНОТОЧНИХ ВИМІРЮВАНЬ

Визначити наближене значення вимірюваноївеличини Х - це значить зробити оцінку математичного сподівання величини Х. При цьому, якщо вимірювана величина Х постійна, то оцінка для $\overline{x}$ є наближене значення істинного значення вимірюваної величини, а якщо вимірювана величина випадкова, то оцінка для $\overline{x}$ є наближене значення математичного очікування вимірюваної випадкової величини.

Необхідність отримання по отриманим даним наближеного значення дисперсії виникає в зв'язку з визначенням характеристики розсіювання вимірюваної випадкової величини.

Нехай є випадкова величина Х з математичним сподіванням $\overline{x}$ і дисперсією D_x; обидва параметри невідомі.Потрібно на підставі отриманих даних знайти спроможні і незміщені оцінки цих параметрів.

Позначимо через X_1, X_2, ..., X_nзначення випадковоївеличини Х, що спостерігаються в результаті проведених nнезалежних рівноточних вимірювань, тобто вимірювань,які проводилися в однакових умовах.

Природно в якості оцінки для математичного сподівання прийняти середнє арифметичне спостережуваних значень

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Покажемо, що ця оцінка є спроможною інезміщенною. Дійсно, відповідно до закону великихчисел, маємо

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_x\right| < \varepsilon\right] = 1$$

Це означає, що $\overline{x}$ є спроможною оцінкою.Оцінка $\overline{x}$ є також і незміщеною, бо

$$M[\overline{x}] = M[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}M[\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = m_x$$

(Спостережувані значення X_1, X_2, ..., X_n розглядаємояк випадкові величини, кожна з якихрозподілена за тим же законом, що і випадкова величина X).

Перейдемо до оцінки дисперсії Dx. Візьмемо статистичну дисперсію і перевіримо її на спроможність і незміщеність.

Статистична дисперсія має вигляд

$$D_x^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{x})^2.$$

Перетворимо цей вираз до іншого виду

$$D_x^* = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\overline{x} + \overline{x^2}) = \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - 2\overline{x}\sum_{i=1}^{n} X_i + n\overline{x^2}\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{x^2}$$

Перший член в правій частині рівності є середнє арифметичне nспостережуваних значень випадкової величини X^2, отже, він сходиться по ймовірності до $M[X^2]$. Другий член $\overline{x^2}$ сходиться по ймовірності до m_x^2. Це означає, що вся права частина рівності сходиться по ймовірності до величини

$$M[X^2] - m_x^2 = D_x$$

Отже, статистична дисперсія D* єспроможною оцінкою дисперсії D_x.

Перевіримо тепер, чи є оцінка D* також і незміщеною. Для цього в формулу для дисперсії замість $\overline{x}$підставимо його вираз і зробимо зазначені дії

$$D_x^* = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{1}{n^2}\sum_{i=1}^{n} X_l^2 - \frac{2}{n^2}\sum_{i<j}^{n} X_i X_j,$$

$$D_x^* = \frac{n-1}{n^2}\sum_{i=1}^{n} X_i^2 - \frac{2}{n^2}\sum_{i<j}^{n} X_i X_j.$$

Так як дисперсія D_x не залежить від того, в якій точці вибрати початок координат, то виберемо його в точці mx і знайдемо математичне сподівання величини.

Маємо

$$M[D_x^*] = \frac{n-1}{n^2}\sum_{i=1}^{n} M[X_i^2] - \frac{2}{n^2}\sum_{i<j}^{n} M[X_i X_j].$$

$$M[D_x^*] = \frac{n-1}{n^2}\sum_{i=1}^{n} D_x - \frac{2}{n}\sum_{i<j}^{n} K_{ij}.$$

З незалежності дослідів випливає, що $K_{ij} = 0$, томуостання рівність приймає вигляд

$$M[D_x^*] = \frac{n-1}{n} D_x$$

Звідси видно, що статистична дисперсія D* є зміщеною оцінкою для дисперсії D_x, її математичне очікування не дорівнює D_x, а кілька меньше.Однако якщо помножити величину D_x на $\frac{n}{n-1}$, то ми отримаємо оцінку для дисперсії D_x, що володіє властивістю незміщеності, бо

$$M\left[\frac{n}{n-1}D_x^*\right] = \frac{n}{n-1}M[D_x^*] = D_x$$

Так як множник $\frac{n}{n-1}$ прямує до одиниці коли $n \to \infty$,то оцінка

$$\widetilde{D} = \frac{n}{n-1}D_x^* = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2$$

буде також і спроможною.

Таким чином, якщо в результаті проведених n незалежних вимірювань випадкової величини X з невідомим математичним сподіванням m_x і дисперсією D_x отримано значення

$$X_1, X_2, ..., X_n$$

то для визначення цих параметрів слід користуватисянаступними наближеними оцінками

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \widetilde{D} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2.$$

8.ВИЗНАЧЕННЯ НАБЛИЖЕНЕГО ЗНАЧЕННЯ ВИМІРЮВАНОЇ ВЕЛИЧИНИ У ВИПАДКУ НЕРІВНОТОЧНИХ ВИМІРЮВАНЬ

Розглянемо визначення наближеного значення математичного очікування m_x деякої величини X понерівноточних вимірам, т. е. за вимірюваннями, кожне з яких характеризується своєю величиною розсіювання.

Нехай ми маємо серію $X_1, X_2, ..., X_n$незалежних вимірювань однієї і тієї ж величини X, дисперсії яких відповідно рівні $\sigma_{x_1}^2, \sigma_{x_2}^2, \ldots, \sigma_{x_n}^2$.

Потрібно за результатами вимірювань знайти оцінку $\overline{x}$, що задовольняє властивостям незміщеності, спроможності та ефективності.

Шукана оцінка $\overline{x}$ є функцією результатів вимірювань, тобто

$$\overline{x} = f\,(X_1, X_2, ..., X_n).$$

Відомо, що найбільш простj. Функціонально. залежністю є лінійна залежність. Тому будемо шукати потрібну нам функцію у вигляді

$$\overline{x} = C_1X_1 + C_2X_2 + \cdots + C_nX_n = \sum_{i=1}^{n} C_iX_i.$$

де $C_1, C_2, ..., C_n$– деякі постійні коефіцієнти, які слід визначити таким чином, щоб оцінка задовольняла умові незміщеності і володіланайменшою дисперсією.

Умова незсуненості виконується, якщо

$$M[\,\overline{x}\,] = M[\sum_{i=1}^{n} C_iX_i] = \sum_{i=1}^{n} C_iM[X_i] = m_x$$

Так як результати вимірювань $X_1, X_2, ..., X_n$, не мають постійної похибки, то

$$M[X_i] = m_x \quad (\text{ i}=1,2,\ldots,\text{n}).$$

Отже, щоб оцінка $\overline{x}$ задовольняла умові незміщеності, необхідним є дотримання рівності

$$\sum_{i=1}^{n} C_i = 1.$$

Тепер будемо вибирати коефіцієнти C_i (i=1,2,…, n) так, щоб дисперсія оцінки була мінімальною, тобто щоб оцінка $\overline{x}$ була ефективною.

Дисперсія оцінки $\overline{x}$, відповідно до властивості дисперсії, дорівнює

$$D[\,\overline{x}\,] = D[\,\sum_{i=1}^{n} C_iX_i] = \sum_{i=1}^{n} C_i^2\, D[X_i]$$

Досліджуємо отриманий вираз на мінімум. Зважаючи на те,що на коефіцієнти C_i вже накладено умова, досліджуємо вираз на умовний мінімум, застосовуючи метод множників Лагранжа.

Складаємо функцію Лагранжа

$$\Phi = \sum_{i=1}^{n} C_i^2 D[X_i] + 2\lambda\left(1 - \sum_{i=1}^{n} C_i\right)$$

Обчислюємо частинні похідні

$$\frac{\partial\Phi}{\partial C_i} = 2C_i\, D[X_i] - 2\lambda \ \text{(i=1,2,…,n).}$$

Прирівнявши праву частину нулю і вирішивши отриманерівняння щодо C_i знаходимо

$$C_i = \frac{\lambda}{D[X_i]} = \frac{\lambda}{\sigma_{x_i}^2}.$$

Позначаючи $\frac{1}{\sigma_{x_i}^2}$ $=g_i$, матимемо

$$C_i = \lambda \cdot g_i.$$

Підставивши знайдене значення C_i, в рівність ,отримаємо

$$\sum_{i=1}^{n} \lambda \cdot g_i = 1.$$

Звідси

$$\lambda = \frac{1}{\sum_{i=1}^{n} g_i}$$

Отже

$$C_i = \lambda \cdot g_i = \frac{g_i}{\sum_{i=1}^{n} g_i}$$

Підставивши отримане значення C_i в формулу,матимемо

$$\overline{x} = \sum_{i=1}^{n} \frac{g_i \cdot x_i}{\sum_{i=1}^{n} g_i}.$$

Перевіримо тепер, чи є оцінка пг також і спроможною. Для цього знайдемо дисперсію величини $\overline{x}$

$$D[\,\overline{x}\,] = D[\sum_{i=1}^{n} \frac{g_i \cdot x_i}{\sum_{i=1}^{n} g_i}] = \frac{1}{(\sum_{i=1}^{n} g_i)^2} \sum_{i=1}^{n} g_i^2 D[X_i],$$

але

$$D[X_i] = \frac{1}{g_i},$$

тому

$$D[\,\overline{x}\,] = \frac{1}{(\sum_{i=1}^{n} g_i)^2} \sum_{i=1}^{n} g_i = \frac{1}{\sum_{i=1}^{n} g_i}$$

Оскільки ряд $\sum_{i=1}^{\infty} g_i$ є розбіжним(не виконується необхідна ознака збіжності ряду), то D $[\overline{x}]$ = 0. Це означає, що функція $\overline{x},$ є спроможноюоцінкою m_x.

Таким чином, оцінка

$$\overline{x} = \frac{\sum_{i=1}^{n} g_i \cdot X_i}{\sum_{i=1}^{n} g_i}$$

для математичного сподівання m_x вимірюваної величини X при нерівноточних вимірах має властивості незсміщеністі, спроможності та ефективності.

Величину

$$g_i = \frac{1}{\sigma_{x_i}^2}$$

прийнято називати вагою і-го виміру.

З формули видно, що чим більше дисперсія $\sigma_{x_i}^2$, тим менше вага g_i результату вимірювання X_i.

9. ДОВІРЧИЙ ІНТЕРВАЛ. ДОВІРЧА ЙМОВІРНІСТЬ

Оцінки, якими ми досі займалися, називаються точковими, так як вони вказують точку на числовій осі, в якій повинно знаходитися значення невідомого параметра. У ряді завдань потрібно не тільки знайти для параметра **а** відповідне числове значення, а й оцінити його точність і надійність. Такого роду завдання дуже важливі при малому числі спостережень, так як точкова оцінка $\tilde{a}$ значною мірою є випадковою і наближена заміна **а** на $\tilde{a}$ може привести до серйозних помилок.

Для визначення точності оцінки $\tilde{a}$–в математичнійстатистиці користуються довірчими інтервалами, а визначення надійності – довірчими йиовірностями. Розкриємо суть цих понять.

Нехай для параметра **а** отримана з експерименту незміщена оцінка $\tilde{a}$. Потрібно оцінити можливу при цьому помилку. Задаємо деяку ймовірність β (наприклад, (β = 0,9) і знаходимо таке значення ε> 0, для якого

$$P(|\tilde{a} - a| < \varepsilon) = \beta.$$

Далі

$$P(\tilde{a} - \varepsilon < a < \tilde{a} + \varepsilon) = \beta.$$

Остання рівність означає, що невідоме значення параметра **а** з ймовірністю βпопаде в інтервал

$$l_\beta(\tilde{a} - \varepsilon; \tilde{a} + \varepsilon).$$

Зауважимо, що тут невідоме значення параметра **а** є невипадковою величиною, а інтервал l_β є випадковою величиною, так як положення інтервалу на осі залежить від випадкової величини $\tilde{a}$ (центр інтервалу), довжина інтервалу 2ε теж в загальному випадку є випадковою величиною. Тому в даному випадку ймовірність β краще тлумачити не як ймовірність попаданняточки **а** в інтервал l_β, а як ймовірність того, що випадковий інтервал l_βнакриє точку.

Інтервал l_βназивається **довірчим інтервалом**, а ймовірність β – **довірчою ймовірністю**. Таким чином, довірчою ймовірністю або надійністюβ, що відповідає даному довірчого інтервалу l_β, називається ймовірність того, що істинне значення параметра **а**належить цьому інтервалу.

Як приклад розглянемо задачу про довірчій інтервал для математичного сподівання.

Нехай проведено n незалежних дослідів над випадковою величиною X з невідомими математичним сподіванням m_x і дисперсією D_x. На підставі даних для цих параметрів побудовані оцінки

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

$$\widetilde{D} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2.$$

Потрібно побудувати довірчий інтервал l_β, відповідний довірчій ймовірності β, для математичного сподівання випадкової величини X. Так як величина $\overline{x}$ являє собою суму n незалежних однаково розподілених випадкових величин X_i, то згідно центральної граничній теоремі її закон розподілу близький до нормального. Користуючись властивостями математичного сподівання і дисперсії, знаходимо

$$M[\overline{x}] = M\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = \frac{1}{n}\sum_{i=1}^{n} m_x = m_x\ ,$$

$$D[\overline{x}] = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} D[X_i] = \frac{1}{n} D_x.$$

Знайдемо тепер таку величину ε_β, для якої

$$P\left(|\overline{x} - m_x| < \varepsilon_\beta\right) = \beta$$

З огляду нате, що закон розподілу випадкової величини$\overline{x}$ близький до нормального, виразимо ймовірність β в лівійчастини рівності через функцію Лапласа

$$P(|\overline{x} - m_x| < \varepsilon_\beta) = \frac{1}{2}\left[\Phi\left(\frac{\varepsilon_\beta}{\sigma_{\overline{x}}\sqrt{2}}\right) - \Phi\left(\frac{-\varepsilon_\beta}{\sigma_{\overline{x}}\sqrt{2}}\right)\right]$$

де $\sigma_{\overline{x}} = \sqrt{\frac{D_x}{n}}$ – середнє і відхилення оцінки.

Так як функція Лапласа непарна, то рівність набуває вигляду

$$P(|\overline{x} - m_x| < \varepsilon_\beta) = \Phi\left(\frac{\varepsilon_\beta}{\sigma_{\overline{x}}\sqrt{2}}\right)$$

З рівняння

$$\Phi\left(\frac{\varepsilon_\beta}{\sigma_{\overline{x}}\sqrt{2}}\right) = \beta$$

знаходимо значення ε_β

$$\varepsilon_\beta = \sigma_{\overline{x}}\sqrt{2}\Phi^{-1}(\beta)$$

де $\Phi^{-1}(\beta)$ –функція, обернена до функції Лапласа.

Величина $\sigma_{\overline{x}} = \sqrt{\frac{D_x}{n}}$, що входить в формулу, виражається через невідому нам дисперсію D_x, тому в якостіїї орієнтовного значення можна взяти оцінку $\widetilde{D}$ і

покласти наближено

$$\sigma_{\overline{x}} \approx \sqrt{\frac{\widetilde{D}}{n}}$$

Таким чином, довірчий інтервал для математичного сподівання наближено дорівнює

$$l_\beta = (\overline{x} - \varepsilon_\beta; \overline{x} + \varepsilon_\beta)$$

де ε_β визначається формулою вище.

10. ПОБУДОВА ДОВІРЧОГО ІНТЕРВАЛУ ДЛЯ МАТЕМАТИЧНОГО СПОДІВАННЯ ВИПАДКОВОЇ ВЕЛИЧИНИ, РОЗПОДІЛЕНОЇ ПО НОРМАЛЬНОМУ ЗАКОНУ. РОЗПОДІЛ СТЬЮДЕНТА

У попередньому параграфі ми розглянули наближений метод побудови довірчого інтервалу для математичного сподівання m_x величини X з невідомимзаконом розподілу. Для точної побудовидовірчого інтервалу необхідно знати закон розподілу випадкової величини

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

який в загальному випадку залежить від самих невідомих параметрів величини X. Виявляється, в деяких випадках від випадкової величини $\overline{x}$ можна перейти до іншої випадкової величини, що є функцією спостережуваних значень X_1, X_2, ..., X_n, закон розподілу якої не залежить від невідомих параметрів величини X, а залежить тільки від числа дослідів і від виду закону розподілу випадкової величини X.

Так, наприклад, доведено, що при нормальномурозподілі величини X випадкова величина

$$T = \sqrt{n}\frac{\overline{x} - m_x}{\sqrt{\widetilde{D}}},$$

де

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

$$\widetilde{D} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2,$$

підпорядковується розподілу Стьюдента з п -1 ступенями свободи. Щільність ймовірності розподілу Стьюдентамає вигляд

$$S_{n-1}(t) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)}\left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}},$$

де $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du$ – гамма-функція.

З формули видно, що розподіл Стьюдента не залежить від $\overline{x}$і$\widetilde{D}$, а залежить тільки від числа дослідів n. При цьому S_{n-1} (t) є парною функцією від t.

Розглянемо застосування розподілу Стьюдента припобудові довірчого інтервалу для математичногосподівання.

Нехай вироблено n незалежних дослідів над випадковою величиною X, розподіленої за нормальнимзакону з невідомими математичним сподіванням m_x ідисперсією D_x. На підставі даних для цихпараметрів побудовані оцінки

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

$$\widetilde{D} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2.$$

Потрібно побудувати довірчий інтервал l_β, відповідний довірчій ймовірності β, для математичного сподівання випадкової величини X. Позначимо через ε_β половину довжини інтервалу, симетричного щодо $\overline{x}$, тоді будемо мати

$$P\left(|\overline{x} - m_x| < \varepsilon_\beta\right) = \beta.$$

Перейдемо в лівій частині рівності від випадкової величини $\overline{x}$ до випадкової величиною T, розподіленої за законом Стьюдента. Для цього помножимо обидві частини нерівності $|\overline{x} - m_x|<\varepsilon_\beta$ на додатню величину $\frac{\sqrt{n}}{\sqrt{\widetilde{D}}}$, отримаємо

$$P\left(\frac{\sqrt{n}|\overline{x} - m_x|}{\sqrt{\widetilde{D}}} < \frac{\varepsilon_\beta}{\sqrt{\frac{\widetilde{D}}{n}}}\right) = \beta,$$

або

$$P\left(T < \frac{\varepsilon_\beta}{\sqrt{\frac{\widetilde{D}}{n}}}\right) = \beta.$$

З огляду на парність функції $S_{n-1}(t)$, отримаємо, що ймовірність β здійснення нерівності

$$T < \frac{\varepsilon_\beta}{\sqrt{\frac{\widetilde{D}}{n}}}$$

дорівнює

$$\mathrm{P}\big(|T| < t_\beta\big) = 2\int_0^{t_\beta} S_{n-1}(t)dt = \beta.$$

Остання рівність визначає величину t_β залежно від довірчої ймовірності β.

Є готова таблиця (див. в додатку), користуючись якою, по довірчій ймовірності β і числу ступенів свободи (n -1) знаходять величину t_β. Визначивши величину t_β, по формулі

$$\varepsilon_\beta = t_\beta \sqrt{\frac{\widetilde{D}}{n}}$$

знаходимо половину ширини довірчого інтервалу l_β. Отже, сам інтервал

$$l_\beta = \left(\overline{x} - t_\beta\sqrt{\frac{\widetilde{D}}{n}}; \overline{x} + t_\beta\sqrt{\frac{\widetilde{D}}{n}}\right).$$

11. ВИЗНАЧЕННЯ НАБЛИЖЕНИХ ЗНАЧЕНЬ ЧИСЛОВИХ ХАРАКТЕРИСТИК СИСТЕМИ ДВОХ ВИПАДКОВИХ ВЕЛИЧИН

Нехай над системою випадкових величин (X, Y) зроблено в однакових умовах n незалежних дослідів.Результати дослідів

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$$

Є незалежними системами випадкових величин, математичні очікування, дисперсіії кореляційні моменти яких однакові, тобто

$$m_{x_i} = m_x,\ \ m_{y_i} = m_y,\ D_{x_i} = D_x, D_{y_i} = D_y, k_{x_i y_i} = k_{xy}.$$

Потрібно шляхом обробки дослідних даних знайти наближені значення зазначених числових характеристик. Це завдання вирішується аналогічно тому, як ми вирішували його для однієї випадкової величина.

Так як невідомі математичні сподівання m_x і m_y, а також дисперсії D_x і D_y є характеристиками окремих випадкових величин, що входять в систему, то для визначення наближених їх значень отримаємо

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

$$\widetilde{D_x} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2, \quad \widetilde{D_y} = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{y})^2.$$

Оскільки кореляційний момент є математичне сподівання добутку відхилень випадкових величин X і Y від своїх математичних сподівань, то наближене значення кореляційного моменту k_{xy} шукаємо як лінійну комбінацію виду

$$\widetilde{k_{xy}} = \sum_{i=1}^{n} C_i (X_i - \overline{x})\,(Y_i - \overline{y}),$$

де C_i –постійні коефіцієнти, причому, в силу рівноточних вимірювань,

$$C_i = C.$$

Невідомий коефіцієнт C визначаємо з умови, щоб величина $\widetilde{k_{xy}}$ була незміщеною оцінкою для кореляційного моменту k_{xy}, тобто щоб

$$M\big[\widetilde{k_{xy}}\big] = C \cdot M\left[\sum_{i=1}^{n}(X_i - \overline{x})\,(Y_i - \overline{y})\right] = k_{xy},$$

Перетворимо вираз, що стоїть під знаком математичного очікування. Так як за умовою m_{x_i} = m_x, то1

$$X_i - \overline{x} = X_i - \frac{1}{n}\sum_{j=1}^{n} X_j + \frac{1}{n}\sum_{j=1}^{n} m_x - \frac{1}{n}\sum_{j=1}^{n} m_x = X_i - m_{x_i} - \frac{1}{n}\sum_{j=1}^{n}\left(X_j - m_{x_j}\right),$$

$$X_i - \overline{x} = \overbrace{X_i}^{0} - \frac{1}{n}\sum_{j=1}^{n}\overbrace{X_j}^{0},$$

аналогічно

$$Y_i - \overline{y} = \overbrace{Y_i}^{0} - \frac{1}{n}\sum_{j=1}^{n}\overbrace{Y_j}^{0}.$$

Отже,

$$M[(X_i - \overline{x})(Y_i - \overline{y})] = M\left[\left(\overbrace{X_i}^{0} - \frac{1}{n}\sum_{j=1}^{n}\overbrace{X_j}^{0}\right)\left(\overbrace{Y_i}^{0} - \frac{1}{n}\sum_{j=1}^{n}\overbrace{Y_j}^{0}\right)\right] =$$

$$= M\left[\overbrace{X_i}^{0}\cdot\overbrace{Y_i}^{0} - \frac{1}{n}\sum_{j=1}^{n}\overbrace{X_i}^{0}\cdot\overbrace{Y_J}^{0} - \frac{1}{n}\sum_{j=1}^{n}\overbrace{X_j}^{0}\cdot\overbrace{Y_i}^{0} + \frac{1}{n^2}\sum_{l=1}^{n}\sum_{j=1}^{n}\overbrace{X_l}^{0}\cdot\overbrace{Y_j}^{0}\right] =$$

$$= k_{x_i y_i} - \frac{1}{n}k_{x_i y_i} - \frac{1}{n}k_{y_i x_i} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} k_{x_i y_j}$$

Беручи до уваги, що $k_{x_i y_i} = k_{xy}$ та $k_{x_i y_j} = 0$, коли $i \neq j$, маємо

$$M[(X_i - \overline{x})(Y_i - \overline{y})] = k_{xy} - \frac{2}{n}k_{xy} + \frac{1}{n}k_{xy} = \frac{n-1}{n}k_{xy}.$$

Таким чином,

$M[\widetilde{k_{xy}}] = C \cdot M[\sum_{i=1}^{n}(X_i - \overline{x})\,(Y_i - \overline{y})] = C \cdot \sum_{i=1}^{n}\frac{n-1}{n}k_{xy} = C(n-1)k_{xy} = k_{xy}$,

якщо $\mathrm{C} = \frac{1}{n-1}$.

Для того, щоб показати спроможність оцінки за умови, що $C_i = C = \frac{1}{n-1}$, знайдемо дисперсію цієї оцінки

$$D\left[\widetilde{k_{xy}}\right] = D\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})(Y_i - \overline{y})\right] = \frac{1}{(n-1)^2}\sum_{i=1}^{n} D[(X_i - \overline{x})(Y_i - \overline{y})]$$

Так як за умовою випадкові величини X_iі Y_i мають однакові розподіли, то D $[Z_i] = D_z$ = const (i = 1, 2, …, n), де $Z_i = (X_i - \overline{x})(Y_i - \overline{y})$.

Отже,

$$D\left[\widetilde{k_{xy}}\right] = \frac{1}{(n-1)^2}\sum_{i=1}^{n} D[Z_i] = \frac{n}{(n-1)^2} D_z.$$

Вираз

$$D\left[\widetilde{k_{xy}}\right] = \frac{n}{(n-1)^2} D_z \to 0, \text{ коли } n\to\infty,$$

а це значить, що

$$\widetilde{k_{xy}} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})\,(Y_i - \overline{y})$$

є незміщеною і спроможною оцінкою для кореляційного моменту k_{xy} системи випадкових величин (X, Y). Ця формула найбільш застосована для визначення кореляційного моменту двох випадкових величин по полученим даним.

Опитний коефіцієнт кореляції $\widetilde{r_{xy}}$ визначають за формулою

$$\widetilde{r_{xy}} = \frac{\widetilde{k_{xy}}}{\widetilde{\sigma_x} \cdot \widetilde{\sigma_y}}$$

При цьому середнє відхилення опитного коефіцієнта кореляції обчислюється за формулою

$$\sigma_{\widetilde{r_{xy}}} = \frac{1-r_{xy}^2}{\sqrt{n-1}} \approx \frac{1-\widetilde{r_{xy}^2}}{\sqrt{n-1}}.$$

12. МЕТОД НАЙБІЛЬШОЇ ПРАВДОПОДІБНОСТІ

Ми розглянули оцінки для математичного сподівання і дисперсії. У цьому параграфі розглянемо один з найважливіших методів для відшукання оцінок параметрів за даними досвіду, який носить назву методу найбільшої правдоподібності.

Нехай функція $f(x,\theta)$, що залежить від параметра θ, є щільністю ймовірності випадкової величини X. Треба на підставі зібраних даних визначити невідомий параметр θ.

Позначимо через $x_1, x_2, ..., x_n$спостережувані значеннявипадкової величини X в результаті проведених n дослідів.

Функцією правдоподібності називається функція

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1,\theta) \cdot f(x_2,\theta) \cdot ... \cdot f(x_n,\theta)$$

Якщо випадкова величина X – дискретна з можливими значеннями

$$\xi_1, \xi_2, \dots, \xi_r,$$

а

$$m_1, m_2, \dots, m_r,$$

дорівнюватимуть відповідно числу спостережених значень, які збігаються з $\xi_1, \xi_2, \dots, \xi_r$, то функціяправдоподібності визначається співвідношенням

$$L(x_1, x_2, \dots, x_n, \theta) = P_1^{m_1}(\theta) \cdot P_2^{m_2}(\theta) \cdot ... \cdot P_r^{m_r}(\theta),$$

де $P_i^{m_i}(\theta) = P(X_i = \xi_i)$, (i=1,2,..,r).

Вважаючи спостережувані значення $x_1, x_2, ..., x_n$ даними, будемо розглядати L як функцію невідомого параметра θ. Сутьметодунайбільшоїправдоподібностіполягаєвтому, що в якості оцінки параметра θ вибирається значення аргументу, який звертає функцію L в максимум. Це значення є функцією від $x_1, x_2, ..., x_n$ і називається оцінкою найбільшої правдоподібності. Звідси згідно відомими правилами диференціального обчислення, для знаходження оцінки найбільшої правдоподібності необхідно вирішити рівняння

$$\frac{\partial L}{\partial \theta} = 0$$

і відібрати те рішення θ, яке звертає функцію L в максимум.

Зазвичай з метою спрощення функцію правдоподібності замінюють її логарифмом і вирішують рівняння

$$\frac{\partial lnL}{\partial \theta} = \frac{1}{L}\frac{\partial L}{\partial \theta} = 0$$

У разі двох параметрів $\theta_1 і \theta_2$ оцінки їх визначаються з двох спільно розв'язуваних рівнянь

$$\frac{\partial L}{\partial \theta_1} = 0\,, \quad \frac{\partial L}{\partial \theta_2} = 0.$$

Метод найбільшої правдоподібності має важливі достоїнства: він завжди призводить до заможних (хоча іноді із зміщеним) оцінками, які мають найменшу можливу дисперсію в порівнянні з іншими і найкращим чином (в певному сенсі) використовують всю інформацію про невідомий параметр, що міститься у вибірці. Однак на практиці він часто призводить до необхідності вирішувати дуже складні системи рівнянь.

13. ЗГЛАДЖУВАННЯ ЕКСПЕРИМЕНТАЛЬНИХ ЗАЛЕЖНОСТЕЙ

При обробці дослідних даних дуже часто доводиться вирішувати завдання, в якій необхідно досліджувати залежність однієї фізичної величини у від іншої фізичної величини х. Наприклад, дослідження залежності величини похибки розміру вироба від температури і т. п.

Нехай проводиться експеримент з метою дослідження залежності величини у від величини х, яка в загальному випадку може бути записана у вигляді

$$y = f(x).$$

Вид цієї залежності і потрібно визначити з експерименту.

Припустимо, що в результаті експерименту отримано ряд точок (x_1, y_1),(x_2, y_2),..., (x_n, y_n), інобудований графік залежності змінної величини у віднезалежної змінної х.

Так як вироблені в ході експерименту вимірювання пов'язані з помилками випадкового характеру, то зазвичай експериментальні точки на графіку мають деякий розкид щодо загальної закономірності. В силу випадковості помилок вимірювання цей розкид або ухилення точок від загальної закономірності також є випадковими. Отже, завдання полягає втакій обробці експериментальних даних, при якій по можливості точно була б відображена тенденція залежності у від х і можливо повніше виключено вплив випадкових, незакономерних ухилень, пов'язаних з похибками екперименту.

Таке завдання є типовою для практики і називається завданням згладжування експериментальної залежності.

Часто буває так, що вид залежності

$$y = f(x)$$

до експерименту відомий з фізичних міркувань, пов'язаних із змістом розв'язуваної задачі, а на підставі експериментальних даних потрібно визначити тільки деякі параметрицієї залежності, які входять в цю залежність лінійно.

При вирішенні задачі згладжування експериментальної залежності в разі, коли вид залежності$y = f(x)$ до експерименту відомий, як правило застосовується розрахунковий метод, відомий під назвою «метод найменших квадратів».

14. МЕТОД НАЙМЕНШИХ КВАДРАТІВ

Метод найменших квадратів застосовується длявирішення різних завдань, пов'язаних з обробкою результатів експерименту. Найбільш важливим застосуванням цього методу є рішення задачі згладжування експериментальної залежності, тобто зображення дослідноїфункціональної залежності аналітичної формулою. При цьому метод найменших квадратів не вирішує питання про вибірзагального вигляду аналітичної функції, а дає можливістьпри заданому типі аналітичної функції $y = f(x)$підібрати найбільш ймовірні значення для параметрів цієїфункції.

Суть методу найменших квадратів при вирішенніпоставленого завдання полягає в наступному.

Нехай отримано n експериментальних точок забсциссами

$$x_1, x_2, \ldots, x_n,$$

і відповідними їм ординатами

$$y_1, y_2, \ldots, y_n.$$

Залежність y від x, зображувана аналітичної функцією $y = f(x)$, не може збігатися з експериментальними значеннями y_iу всіх n точках. Це означає, що для всіх або деяких точок різниця

$$\Delta_i = y_i - f(x_i)$$

буде відмінна від нуля.

Потрібно підібрати параметри функції таким чином, щоб сума квадратів різниць була найменшою, тобто потрібно звести до мінімума вираз

$$z = \sum_{i=1}^{n} \Delta_i^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2.$$

Таким чином, при методі найменших квадратівнаближення аналітичної функції $y = f(x)$доекспериментальної залежності вважається найкращим, якщо

виконується умова мінімуму суми квадратів відхилень шуканої аналітичної функції від експериментальної залежності.

Слід зауважити, що записане вираз являєсобою поліном другого ступеня щодо невідомихпараметрів (невідомі параметри в залежність $y = f(x)$входять лінійно), який не може приймативід'ємних значень. Тому існують такі значення невідомих параметрів, при яких функція досягає мінімуму, і цей мінімум в залежності від значень x_i і y_i буде додатним або рівним нулю.

При вирішенні багатьох практичних завдань функціональну залежність у від х шукають у вигляді

$$y = \sum_{k=1}^{m} a_k f_k(x),$$

де $f_1(x), f_2(x), \dots, f_m(x)$ – відомі функції, a_1, a_2,…, a_m – невідомі параметри.

Так, наприклад, при дослідженні коливальних процесів функціями $f_k(x)(k = 1,2, \dots, m)$ є тригонометричні функції

$$f_k(x) = coskx, f_k(x) = sinkx.$$

При дослідженні в багатьох областях дужечасто зустрічаються степеневі функції

$$f_k(x) = x^{k-1}(k = 1,2, \dots, m).$$

Таким чином, $f_k(x)$ є відомими елементарними функціями аргументу х.

Виходячи з принципу найменших квадратів, необхідно підібрати такі значення невідомих параметрівa_1, a_2,…, a_m, при яких звертається в мінімум вираз

$$z=\sum_{i=1}^{n}(y_i - \sum_{k=1}^{m} a_k\, f_k(x_i))^2.$$

Записаний вираз є функцією невідомих параметрів a_k, тому для відшукання мінімуму цієї функції потрібно згідно з правилами диференціального

обчислення знайти частинні похідні функції z повсім параметрам a_k (k = 1,2,…, m) і приравняти їх нулю

$$\frac{\partial z}{\partial a_1} = 2\sum_{i=1}^{n}\left(y_i - \sum_{k=1}^{m} a_k\, f_k(x_i)\right)\cdot\left(-f_1(x_i)\right) = 0$$

$$\frac{\partial z}{\partial a_2} = 2\sum_{i=1}^{n}\left(y_i - \sum_{k=1}^{m} a_k\, f_k(x_i)\right)\cdot\left(-f_2(x_i)\right) = 0$$

…………………………………………………...

$$\frac{\partial z}{\partial a_m} = 2\sum_{i=1}^{n}\left(y_i - \sum_{k=1}^{m} a_k\, f_k(x_i)\right)\cdot\left(-f_m(x_i)\right) = 0$$

Підставляючи в систему відомі значення x_i, y_i отримаємо систему m лінійних рівнянь щодо невідомих параметрів a_k, рішення якої може бути отримано за допомогою формул Крамера або методом Гаусса.

Розглянемо застосування методу найменших квадратів,коли для зображення експериментальної залежностіобрана парабола другого порядку

$$y = ax^2 + bx + c.$$

Нехай в результаті незалежних експериментів отримано n значень величини y

$$y_1, y_2, \ldots, y_n,$$

які відповідають значенням величини x

$$x_1, x_2, \ldots, x_n.$$

Для визначення невідомих параметрів a, b і c методом найменших квадратів складаємо суму квадратів відхилень шуканої аналітичної функції від спостережуваних значень в даних точках

$$z = \sum_{i=1}^{n}\left(y_i - ax_i^2 - bx_i - c\right)^2$$

Диференціюючи цю функцію по невідомим параметрама, b і с і прирівнюючи похідні до нуля, отримаємо наступну систему рівнянь

$$-\frac{1}{2}\frac{\partial z}{\partial a} = \sum_{i=1}^{n}(y_i - ax_i^2 – bx_i - c)x_i^2 = 0$$
$$-\frac{1}{2}\frac{\partial z}{\partial b} = \sum_{i=1}^{n}(y_i - ax_i^2 – bx_i - c)x_i = 0$$
$$-\frac{1}{2}\frac{\partial z}{\partial c} = \sum_{i=1}^{n}(y_i - ax_i^2 – bx_i - c) = 0$$

або перетворивши рівняння, отримаємосистему рівнянь:

$$\mathrm{a}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^4 + \mathrm{b}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^3 + \mathrm{c}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^2 = \sum_{i=1}^{n} y_i x_i^2$$

$$\mathrm{a}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^3 + \mathrm{b}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^2 + \mathrm{c}\sum_{\mathrm{i}=1}^{\mathrm{n}} x_i = \sum_{i=1}^{n} y_i x_i$$

$$\mathrm{a}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^2 + \mathrm{b}\sum_{\mathrm{i}=1}^{\mathrm{n}} x_i + \mathrm{nc} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{y_i}$$

Отримана система являє собою систему трьохлінійних рівнянь щодо невідомих параметріва, b і с. Вирішуючи систему за допомогою формул Крамера або методом Гаусса, отримаємо значення параметрів a, b і сза методом найменших квадратів.

15. СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ

Перш ніж формулювати завдання перевірки гіпотезв загальному вигляді, розглянемо два приклади.

Приклад 1. Є склад готової продукції. Відомо, що вироби надходять на склад партіями з двох заводів, що випускають продукцію різної якості, і такими ж партіями відпускаються споживачеві. Якість продукції заводу характеризується ймовірністю P того, що навмання обраний виріб є бракованим. Для одного заводаP=P0, для іншого $P=P_1$ ($P_0>P_1$). Споживач навмання вибирає одну партію виробів. Потрібно на підставі результатів контролю вирішити, на якому заводі виготовлена обрана партія виробів.

Р і ш е н н я. H_0– гіпотеза, яка полягає в тому, що обрана партія виробів має погану якість, тобто ймовірність браку дорівнює P_0; H_1 – протилежна гіпотеза, ймовірність браку дорівнює P_1. Будемо називати H_0 – нульовою гіпотезою, а

H_1 – конкуруючою гіпотезою.

Відберемо з партії навмання n виробів. Нехай Y позначає кількість бракованих виробів серед відібраних. Ясно, що Y є випадковою величиною, можливими значеннями якої будуть 0, 1, 2, ..., n. Під рішенням поставленого завдання розуміється вироблення вирішального правила, яке зіставляє кожному можливому значенню випадкової величини Y одну з гіпотез H_0 або H_1.Позначімо набір можливих значень випадкової величини Y через Δ, тоді відповідно до сказаного вище, шукане вирішальне правило полягає в деякому розбитті множини Δ на частини Δ_0 і Δ_1. При попаданні можливого значення випадкової величини Y в множину Δ_0 приймається гіпотеза H_0 і, навпаки, при попаданні можливого значення в множинуΔ_1 приймається гіпотеза H_1.

Питання полягає в тому, яке з можливихрозбиття множини Δ на частини Δ_1 і Δ_2 слід вибрати.

Загальна постановка задачі. Є дві протилежні гіпотези H_0 іH_1 і деяка пов'язана з ними випадкова величина Y. Нехай y позначає числове значення випадкової величини Y, отримане в результаті випробування, $\Delta -$ множина всіх можливих значень випадкової величини Y. Потрібно провести перевірку нульової гіпотези H_0щодо конкуруючої гіпотези H_1 на підставі результатів випробування.

Розіб'ємо множинуΔ на дві частини Δ_0 і Δ_1 з умовою прийняття гіпотези H_0 при попаданні отриманого значення y випадкової величини Y в результаті проведеного експерименту в Δ_0 і гіпотези H_1 - при попаданні y в Δ_1. Вибір вирішального правила, тобто правила розбиття множини Δ на дві частини Δ_0 і Δ_1в будь-якій задачі перевірки гіпотез можливий більше, ніж одним способом.

Так, можна задати будь-яке число l_1і покласти

$$\Delta_0 = (-\infty; l_1), \Delta_1 = (l_1; +\infty),$$

Або

$$\Delta_0 = (-\infty; l_1) \cup (l_2; +\infty), \Delta_1 = (l_1; l_2).$$

Питання: якому з цих розбиттів слід віддати перевагу, або, відволікаючись від розглянутого прикладу, яке з усіх можливих розбиттів в кожній конкретній задачі вважати найкращим?

Метод мінімуму ризику. Для застосування до задачі перевірки гіпотез методу мінімуму ризику потрібно мати у своєму розпорядженні деякими імовірнісні дані. Будемо вважати відомими два умовних розподілу ймовірностей випадкової величини Y:

$f_0(y)$ – при умови, що вірна гіпотеза H_0;

$f_1(y)$ – при умови, що вірна гіпотеза H_1.

Зауважимо, що випадкова величина Y може бути і дискретної і неперервної. Тоді під функціями $f_0(y)$ і $f_1(y)$ в першому випадку будемо розуміти умовні дискретні розподілу ймовірностей, в другому випадку умовні щільності розподілу.

У прикладі ймовірність P того, що навмання обраний виріб є бракованим не залежить від результатів перевірки інших виробів і за умови істинності гіпотези H_0, дорівнює P_0. Тому закон розподілу випадкової величини Y є біноміальним і, отже, має вигляд

$$f_0(y) = P(Y = y | \, P = P_0) = C_n^y P_0^y (1 - P_0)^{n-y}.$$

Аналогічно, за умови істинності гіпотези H_1

$$f_1(y) = P(Y = y|\ P = P_1) = C_n^y P_1^y (1 - P_1)^{n-y}.$$

Крім умовних розподілів $f_0(y)$ і $f_1(y)$ потрібно знати апріорну ймовірність Р того, що гіпотеза H_0має місце.

Іноді є відомості про цю ймовірность, а іноді нічого не відомо чи відомо дуже мало. Так, наприклад, якщо в задачі про приймальний контроль відомо, що серед партій готової продукції, що зберігається на складі, одна чверть поганої якості, то за умови, що споживач навмання вибирає одну партію виробів,

$$\mathrm{P} = \frac{1}{4}.$$

Будемо розглядати такі випадкові події:

A –вірна гіпотеза H_0,

$\overline{A}$ – вірна гіпотеза H_1,

B– результат експерименту y потрапив в область Δ_0,

$\overline{B}$– результат експерименту y потрапив в область Δ_1.

Тоді в результаті прийняття рішення можливий один з наступних чотирьох випадків:

AB – вірна гіпотеза H_0 і прийнято рішення про її істинность;

$\overline{A}B$– вірна гіпотеза H_1, а прийнято рішення про істинность гіпотези H_0;

$A\,\overline{B}$ – вірна гіпотеза H_0, а прийнято рішення про істинності гіпотези H_1;

$\overline{AB}$ – вірна гіпотеза H_1 і прийнято рішення про її істинность.

Звідси видно, що результати $\overline{A}B$ і $A\,\overline{B}$ пов'язані з помилковими рішеннями. Результату $\overline{A}B$ відповідає так звана помилка **першого роду**, а результату $A\,\overline{B}$ -помилка**другого роду**.

Наприклад, в задачі виявлення бракованих виробів помилці першого роду відповідає прийняття рішення про наявність браку в разі його відсутності, помилку другого родувідповідає прийняття рішення про відсутність браку в разі його наявності.

Для відповіді на питання, яке з вирішальних правил слід вважати найкращим, введемо поняття функції втрат і середнього ризику.

Функція втрат зіставляє кожному з чотирьохможливих результатів $AB, \overline{A}B, A\,\overline{B}, \overline{AB}$відповідні втрати, виражені в деякій системі одиниць.

При правильному вирішенні природно покласти втратирівними нулю. Втрати, пов'язані з помилками першогоі другого роду, позначимо відповідно C_1 і C_2. Будемо вважати, що C_1 і C_2 – додатні. Для завдання функції втрат потрібно вказати ці два числа. Надалі передбачається, що функція втрат задана.

Зауважимо, однак, що в практичних завданнях часто важко зробити обґрунтований вибір величини C_1 і C_2.

Нижче ми покажемо, як надходять в таких випадках.

Перейдемо тепер до поняття ризику. Нехай P_0, P_1 іP_2 –ймовірності відповідно правильного рішення,помилки першого роду і помилки другого роду.

Визначення значень цих ймовірностей буде приведено нижче.

Величина втрат С, до яких призводить одноразове застосування вирішального правила, є випадковоювеличиною, що приймає значення О, C_1, C_2зймовірностями відповідно P_0, P_1, P_2.

Математичне сподівання М [С] випадкової величини С називається середнім ризиком (або просто ризиком) і позначається буквою r.

Таким чином,

$$r = \mathrm{M}\,[\mathrm{C}] = P_0 \cdot 0 + P_1 \cdot C_1 + P_2 \cdot C_2 = P_1 \cdot C_1 + P_2 \cdot C_2$$

Поняття ризику призводить до природного способупорівняння вирішальних правил. З двох правил кращимвважається те, яке призводить до меншого ризику.

Оптимальним вирішальним правилом називаєтьсяправило, яке приводить до найменьшого можливого в данному завданні ризику.

Отже, треба знайти оптимальне вирішальнеправило, яке відповідає заданим умовним розподілам$f_0(y)$ і $f_1(y)$, апріорної ймовірності Р і функції втрат (C_1, C_2). Це правило будемо позначатибуквою Г. При знаходженні правила Г обмежимосярозглядом неперервної випадкової величини Y.

Позначимо через α умовну ймовірність помилки першого роду, обчислену за умови істинності гіпотези H_0, а через β – умовну ймовірність помилки другого роду, обчислену за умови істинності гіпотези H_1. Застосовуючи правило визначення ймовірності попадання випадкової величини Y на задану ділянку, якщо відома її щільність ймовірності, запишемо ймовірності α іβ за допомогою умовних розподілів $f_0(y)$ і $f_1(y)$

$$\alpha = P\left(\overline{B}|\,A\right) = \underbrace{\int f_0(y)dy}_{\Delta_1}, \beta = P\left(B|\,\overline{A}\right) = \underbrace{\int f_1(y)dy}_{\Delta_0}.$$

У свою чергу, безумовні ймовірності P_1 і P_2 помилок першого і другого роду, виражаються через умовні ймовірності цих помилок α й β і апріорну ймовірність P наступним чином

$$P_1 = P\left(A\,\overline{B}\right) = P(A)P\left(\overline{B}|\,A\right) = P \cdot \alpha,$$

$$P_2 = P\left(\overline{A}\,B\right) = P\left(\overline{A}\right)P\left(B|\,\overline{A}\right) = (1 - P) \cdot \beta.$$

Підставимо отримані значення P_1 і P_2 в формулу для ризику

$$r = P \cdot \alpha \cdot C_1 + (1 - P) \cdot \beta \cdot C_2$$

$$r = P \cdot \underbrace{\int f_0(y)dy}_{\Delta_1} \cdot C_1 + (1 - P) \cdot \underbrace{\int f_1(y)dy}_{\Delta_0} \cdot C_2$$

З формули видно, що кожному способу розбиття множини Δ на області Δ_0 і Δ_1 відповідає своє значення ризику. Потрібно вибрати області Δ_0 і Δ_1 так, щоб вираз для ризику досяг мінімуму.

Використовуючи властивості щільності ймовірності

$$\underbrace{\int f_0(y)dy}_{\Delta} = \underbrace{\int f_0(y)dy}_{\Delta_0} + \underbrace{\int f_0(y)dy}_{\Delta_1}=1$$

перепишемо формулу для ризику у вигляді

$$r = P \cdot (1 - \underbrace{\int f_0(y)dy}_{\Delta_0}) \cdot C_1 + (1 - P) \cdot \underbrace{\int f_1(y)dy}_{\Delta_0} \cdot C_2,$$

$$r = P \cdot C_1 + \underbrace{\int [(1 - P)}_{\Delta_0} \cdot C_2 \cdot f_1(y) - P \cdot C_1 \cdot f_0(y)]dy.$$

Останній вираз досягає мінімуму при такому виборі області Δ_0, який призводить до найменшого можливого значення інтеграла в правій частині. А для того, щоб інтеграл був мінімальним, потрібно включити до складу Δ_0 ті і тільки ті значення y, в яких підінтегральна функція від'ємна, тобто

$$(1 - P) \cdot C_2 \cdot f_1(y) - P \cdot C_1 \cdot f_0(y) < 0,$$

а до складу Δ_1 - інші значення y.

Запишемо нерівність в такому вигляді

$$\frac{f_1(y)}{f_0(y)} < \frac{PC_1}{(1-P)C_2}.$$

Функція $\frac{f_1(y)}{f_0(y)}$ називається **відношенням правдоподібності**.

Таким чином, шукане оптимальне вирішальне правило Г полягає в наступному: для отриманого в результаті експерименту значення y обчислюється відношення правдоподібності $\frac{f_1(y)}{f_0(y)}$ і порівнюється з числом

$$l = \frac{PC_1}{(1-P)C_2};$$

якщо відношення правдоподібності менше l, приймається гіпотеза H_0; в іншому випадку - гіпотеза H_1.

Оптимальне вирішальне правило Г називається **пороговим критерієм** для відношення правдоподібності $\frac{f_1(y)}{f_0(y)}$ з порогом $l = \frac{PC_1}{(1-P)C_2}$.

Аналогічний результат виходить і для дискретної випадкової величини.

Застосуємо пороговий критерій Г до розглянутого вище прикладу.

Приклад 1. Використовуючи рівності, наведені вище,знайдемо відношення правдоподібності

$$\frac{f_1(y)}{f_0(y)} = \frac{C_n^y P_1^y (1-P_1)^{n-y}}{C_n^y P_0^y (1-P_0)^{n-y}} = \frac{[P_1(1-P_0)]^y}{[P_0(1-P_1)]^y} \left(\frac{1-P_1}{1-P_0}\right)^n$$

Отже, нерівність для даного прикладу набирає вигляду

$$\frac{C_n^y P_1^y (1-P_1)^{n-y}}{C_n^y P_0^y (1-P_0)^{n-y}} = \frac{[P_1(1-P_0)]^y}{[P_0(1-P_1)]^y} \left(\frac{1-P_1}{1-P_0}\right)^n < \frac{PC_1}{(1-P)C_2},$$

звідки

$$\frac{[P_1(1-P_0)]^y}{[P_0(1-P_1)]^y} < \frac{PC_1}{(1-P)C_2} \left(\frac{1-P_0}{1-P_1}\right)^n,$$

З умови, що $P_0 > P_1$, слідує нерівність

$$\frac{P_1(1-P_0)}{P_0(1-P_1)} < 1,$$

а це значить, що

$$ln\frac{P_1(1-P_0)}{P_0(1-P_1)} < 0.$$

Тому, визначаючи y з нерівності, будемо мати

$$y > \frac{\ln[\frac{PC_1}{(1-P)C_2}\left(\frac{1-P_0}{1-P_1}\right)^n]}{\ln[\frac{P_1(1-P_0)}{P_0(1-P_1)}]}.$$

Отже, якщо число y бракованих виробів серед навмання обраних n виробів задовольняє нерівності, то приймається рішення про погану якість отриманої партії, в іншому випадку– рішення про хорошу якість.

16. ПОНЯТТЯ ПРО КРИТЕРІЙ ЗГОДИ

У багатьох випадках практики на підставі тих чи інших даних робиться припущення про вид закону розподілу, що цікавить нас випадкової величини X.

Однак для остаточного вирішення питання про вид закону розподілу в подібних випадках видається доцільним перевірити, наскільки зроблене припущення узгоджується з досвідом. При цьому з огляду на обмежене число спостережень закон розподілу зазвичай буде в якійсь мірі відрізнятися відпередбачуваного, навіть якщо припущення про законрозподілу зроблено правильно. У зв'язку з цимвиникає необхідність вирішувати таку задачу:чи є розбіжність між опитним закономрозподілу і передбачуваним законом розподілунаслідком обмеженого числа спостережень, або воноє істотним і пов'язане з тим, щодійсний розподіл випадкової величини відрізняєтьсявід передбачуваного. Для вирішення поставленого завданняслужать так звані «критерії згоди».

Ідея застосування критеріїв згоди полягає внаступному.

Нехай, наприклад, на підставі даногостатистичного матеріалу нам належить перевірити гіпотезу H,яка полягає в тому, що випадкова величина X маєфункцію розподілу F (x).Для того щоб прийняти або спростуватигіпотезу H, будемо розглядати випадкову величину Y, яка характеризує ступінь розбіжності теоретичногоі статистичного розподілів. Величину Y можнавибирати різними способами.

Наприклад, в якості Y можна взяти максимальне відхилення статистичноїфункції розподілу F* (x) від теоретичної функції розподілу F (x).Очевидно, закон розподілу випадкової величини Y залежить від закону розподілу випадкової величини X, над якою проводилися досліди, і від числа дослідів n. Припустимо, що закон розподілу випадкової величини нам відомий.

Нехай в результаті проведених n дослідів над випадковою величиною X величина Y прийняла деяке значення y. Питається, чи можна пояснити прийняте значення Y = у випадковими причинами або ж це значення занадто велике і вказує на наявність суттєвої різниці між теоретичним і статистичним розподілами, тобто непридатність гіпотези H?Для відповіді на це питання припустимо, що вірна гіпотеза Н і обчислимо ймовірність того, що випадкова величина Y за рахунок випадкових причин, пов'язаних з обмеженим обсягом досвідченого матеріалу, прийме значення не менше, аніж спостережене значення у, тобто обчислимо ймовірність Р (Y ≥ у). Якщо ця ймовірність мала, то гіпотезу Н слід спростувати як малоправдоподібну, а якщо ж ця ймовірність значна, то експериментальні дані не суперечатьгіпотезі Н.

Для обчислення ймовірності Р {У ≥ у) необхідно знати закон розподілу випадкової величини Y, який, як ми вже відзначали, залежить від закону розподілу випадкової величини X (функції розподілу F (x)) і від числа дослідів n. Виявляється, що при деякі способи вибору випадкової величини Y її закон розподілу при досить великому n практично не залежить від закону розподілу випадкової величини X. Саме такими заходами розбіжності і користуються в математичній статистиці в якості критеріїв згоди.

Найбільш простим критерієм перевірки гіпотези про вид закону розподілу є критерій академіка О.Н.Колмогорова, що представляє собою максимальне значення абсолютної величини різниці між статистичною функцією розподілу F*(x) і відповідною теоретичною функцією розподілу F(x), тобто

$$D = max|F^{*}(x) - F(x)|.$$

О.Н.Колмогоров довів, що який би вид не мала безперервна функція розподілу F(x) при необмеженому зростанні числа незалежних спостережень n, вірогідність нерівності

$$D\sqrt{n} \geq \lambda$$

прямує до межі

$$P(\lambda) = 1 - \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2}.$$

Для ймовірності P (X) складена таблиця, коротка витримка з якої наводиться нижче

λ	0.0	0.5	1.0	1.5	2.0
$P(\lambda)$	1.000	0.964	0.270	0.022	0.001

Схема застосування критерію О.Н.Колмогорова наступна.

1. За результатами n проведених вимірювань будується статистична функція розподілу F * (x).

2. На тому ж графіку будується передбачувана теоретична функція розподілу F (x).

3. Визначається максимальна величина модулярізниці їх ординат.

4. Обчислюється величина

$$\lambda = D\sqrt{n}$$

5. По таблиці (частина якої приведена вище) знаходиться ймовірність P(λ), що відповідає тому, що за рахунок випадкових причин максимальна розбіжність між F*(x) і F (x) буде не менше, ніж фактично спостерігається. Якщо ймовірність P(λ) дуже мала, то гіпотеза бракується: при порівняно великій ймовірністі P(λ) гіпотеза вважається сумісною з результатами досвіду.

Зауважимо, що критерій О.Н.Колмогорова може застосовуватися тільки в разі, коли гіпотетичний розподіл F {x) повністю відомий, тобто відомий не тільки вид функції розподілу F (x), а й всі параметри, що в неї входять. Очевидно, що такі випадки на практиці зустрічаються рідко. Зазвичай з теоретичних міркувань відомий тільки вид функції P(x), параметри її доводиться визначати за результатами вибірки. У таких випадках слід застосовувати інші критерії згоди. Один з найбільш часто вживаних на практиці критеріїв згоди, який дозволяє проводити перевірку гіпотези відповідності досвідченого закону розподілу передбачуваного (теоретичному) не тільки у випадках, коли останній відомий повністю, але і тоді, коли параметри передбачуваного закону розподілу визначаються на підставі досвідчених даних, є критерій Пірсона χ^2(хі-квадрат).

ТИПОВІ РОЗРАХУНКИ

До умови задачі додається вибірка відповідно до кожного з варіантів.

Після варіантів вибірок подається приклад виконання 0-го варіанта.

Для даної вибірки (див. відповідний варіант після умови задачі) необхідно :

1. Скласти варіаційний ряд, впорядкувавши за зростанням числові дані (варіанти) вибірки.
2. Скласти статистичний ряд, приписавши кожній з варіант відповідні їм частоти.

Для статистичного ряду знайти:

а) середнє арифметичне,

б) незміщені дисперсію і стандартне відхилення.

3. Побудувати:

а) полігон відносних частот статистичного ряду,

б) гістограму відносних частот,

в) функцію накопичених відносних частот та її графік.

4. Знайти довірчі інтервали для математичного сподівання при довірчій ймовірності γ

$$\gamma = \begin{cases} 0.8, \text{якщо } V \leq 10 \\ 0.95, \text{якщо } 10 < V \leq 20 \\ 0.98, \text{якщо } 20 < V \end{cases},$$

де V- номер варіанту.

5. При рівні значущості α перевірити гіпотезу про нормальний розподіл відповідної генеральної сукупності, користуючись критерієм згоди Колмогорова, якщо

$$\alpha = \begin{cases} 0.1, \text{якщо } V \leq 10 \\ 0.05, \text{якщо } 10 < V \leq 20 \\ 0.02, \text{якщо } 20 < V \end{cases}$$

Далі подаються варіанти вибірок

Варіант 1.

Дані про продаж чоловічих зимових черевиків за розмірами

42	39	38	42	42	41	42	39	38	42	42	41
44	41	40	40	39	42	44	41	40	40	39	42
42	40	43	41	42	40	42	40	43	41	42	40
41	38	39	38	43	41	41	38	39	38	43	41
43	43	41	44	42	42	43	43	41	44	42	42
42	41	40	42	44	43	42	41	40	42	44	43
42	37	42	40	41	39	42	37	42	40	41	39
43	42	43	42	43	41	43	42	43	42	43	41
44	41	41	41	40	42	44	41	41	41	40	42
42	42	41	42	40	42	42	42	41	42	40	42

Варіант 2.

Зареєстрована продаж таких розмірів дитячого взуття.

33	34	33	33	34	32	33	32	33	34	33	33	34	33
32	31	34	33	33	34	33	32	31	34	31	33	34	33
34	32	31	34	32	34	32	34	34	32	31	34	32	32
31	34	33	33	34	34	34	34	31	34	33	33	34	34
32	33	32	31	33	34	33	34	34	33	34	31	33	33
34	34	32	33	34	32	31	32	34	34	32	33	34	31
32	31	33	34	33	33	34	33	32	31	33	34	33	34
34	32	34	33	32	34	33	34	34	32	34	33	32	33
33	34	34	34	34	33	34	33	33	34	34	34	34	31
32	31	33	32	33	31	32	31	32	31	33	32	33	32

Варіант 3.

Дані продажу розмірів чоловічих сорочок.

52	50	52	52	48	48	52	50	52	52	48	48	52	50
50	52	50	48	52	50	50	52	50	48	52	50	52	48
54	48	48	50	50	52	54	48	48	50	50	52	48	52
56	46	50	46	48	48	56	46	50	46	48	48	50	52
50	50	54	50	56	48	50	50	54	50	56	48	50	48
56	48	48	44	50	46	56	48	48	44	50	46	52	50
54	48	50	40	46	48	54	48	50	50	48	48	48	50
50	48	48	50	48	48	50	48	48	50	48	48	48	52
48	54	48	54	48	48	50	49						

Варіант 4.

Оцінки, отримані студентами 2-го курсу під час сесії.

3, 3, 4, 4, 4, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3, 5, 4, 3, 3, 4, 4, 4, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3, 3, 2,5, 2, 4, 4, 4, 5, 3, 2, 2, 4, 4, 4, 5, 5, 5, 4, 3, 3, 4, 4, 4, 5, 4, 3, 2, 5, 4, 4, 4, 4, 4, 5, 3, 2, 2, 4, 4, 4, 5, 5, 5, 4, 3, 3, 4, 4, 4, 5, 4, 3, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3, 4, 4, 5, 5, 5, 4, 3, 3, 4, 4, 4, 5, 4, 3, 5, 4, 4, 5, 5, 5, 4, 3, 3, 4, 4, 4, 5, 4, 3.

Варіант 5.

Дані про продаж жіночих чобіток за розмірами.

35	36	37	35	35	36	35	36	37	35	35	36	37	38
37	37	36	36	37	40	36	36	39	37	36	39	37	39
37	37	36	36	37	40	36	36	39	37	36	39	37	34
38	38	37	36	38	38	38	38	37	36	38	38	35	36
36	36	36	37	36	40	36	36	36	37	36	40	38	37
37	33	38	36	37	38	37	33	38	36	37	38	37	35
36	38	36	34	35	39	36	38	36	34	35	39	37	38
34	35	35	38	36	37	34	35	35	38	36	37	35	37
36	39	37	35	34	38	37	37						

Варіант 6.

Дані про продаж чоловічих костюмів за розмірами.

50	50	50	52	48	48	52	50	52	52	48	48	52	48
50	52	50	48	52	50	50	52	50	48	52	50	54	54
54	48	48	50	50	52	54	48	48	50	50	52	48	50
56	46	50	46	46	46	56	46	50	44	48	48	52	48
50	50	54	50	56	48	50	50	44	48	56	50	50	52
56	48	48	44	50	46	56	48	48	44	50	46	52	52
54	48	50	50	48	48	54	48	50	50	48	48	44	48
50	48	48	50	48	48	50	48	48	50	48	46	48	50
52	48	54	48	50	50	48	48						

Варіант 7.

Розряди виробників на виробництві

7	6	7	5	3	6	3	6	5	6	5
7	6	7	5	3	7	6	3	5	6	7
6	7	6	5	3	7	7	3	5	4	6
6	5	6	7	4	6	7	4	4	4	6
5	5	6	7	5	5	5	4	3	4	4
4	4	5	7	6	4	5	4	7	3	4
3	4	4	4	7	3	4	6	7	3	3
6	4	4	4	4	6	4	6	6	4	3
6	6	4	3	5	6	3	6	5	4	4

Варіант 8.

Сім’я за кількістю дітей.

1	5	2	0	4	0	2	5	3	2	6	0
1	2	2	0	5	0	2	2	2	2	2	0
0	2	3	1	2	0	3	3	1	0	4	0
0	3	4	1	2	2	4	1	2	0	5	4
2	1	2	1	3	1	5	0	2	2	4	3
2	1	2	2	3	1	7	0	1	2	1	2
3	1	1	2	1	1	1	2	1	1	1	2
3	0	1	2	1	1	2	1	3	1	1	2
4	0	1	2	0	0	3	1	3	1	2	2
4	1	0	2	0	0	4	1	2	3	0	1

Варіант 9.

Зареєстрований продаж розмірів дитячого взуття.

32	32	34	32	34	32	33	32	33	34	33	33	34	33
32	31	34	31	33	33	34	33	32	31	34	31	33	34
32	32	31	34	32	34	32	34	34	32	31	34	32	32
31	34	33	33	34	34	34	34	31	34	30	33	34	34
32	33	34	31	33	34	33	30	32	33	31	33	34	33
34	34	32	33	30	32	31	32	34	34	32	33	30	31
32	31	33	34	33	33	34	33	32	31	33	34	33	34
34	32	34	33	32	34	33	34	34	32	34	33	32	33
33	34	34	34	34	33	34	33	33	34	34	34	34	34
32	30	33	32	33	31	32	31	32	31	33	32	33	32

Варіант 10.

Кількість викликів на станцію швидкої допомоги за добу.

35	36	37	35	35	36	35	36	37	35	36	36	31
37	37	36	36	37	40	37	37	36	36	37	40	33
36	36	39	37	36	36	39	37	36	39	36	39	32
38	38	37	36	38	38	38	38	37	36	38	33	33
36	36	36	37	36	40	36	36	36	37	36	40	32
31	33	38	36	37	38	37	33	38	36	37	38	
36	38	36	34	35	39	36	38	36	34	35	39	
34	35	35	38	36	37	34	35	35	38	36	37	
37	37	38	37	34	38	37	37	38	37	34	38	
37	35	37	35	36	39	37	35	37	35	35	36	
32	33	34	33	33	34	33	33	32	31	34	32	
32	34	31	34	33	33	34	34	34	32	33	34	
31	33	34	32	32	34	33	34	34	32	34	33	
32	33	33	33	34	34	34	34	34	34	31	32	

Варіант 11.

Дані про продаж чоловічих полуботок за розмірами.

44	42	46	42	42	41	42	39	38	42	42	41	41	39
44	41	40	40	39	42	44	41	45	40	39	42	43	40
42	40	43	41	42	40	46	40	43	41	42	46	41	43
41	38	39	38	43	41	41	38	39	43	43	41	42	41
43	43	41	44	42	42	43	43	41	46	42	42	43	38
42	41	40	42	44	43	41	42	45	42	44	43	44	38
42	37	42	40	41	39	42	37	42	40	41	39	43	41
43	42	43	42	43	41	43	42	43	42	43	41	40	42
44	41	41	41	40	42	44	44	41	41	40	42	42	42
42	42	41	42	40	42	42	42	41	42	42	40	42	43
42	39	38	42	42	41	44	41	40	40	39	42	42	42
44	41	41	41	40	42	42	41	39	42	40	41	42	42

Варіант 12.

Число обривів нитки на ткацьких станках у цеху за 5 хв.

32	33	34	33	33	34	33	35	36	37	35	35	36
33	32	31	34	31	33	34	37	37	36	36	37	40
34	34	32	31	34	32	32	36	36	39	37	36	39
34	31	34	33	33	34	34	38	38	37	36	38	38
34	32	33	31	33	34	33	36	36	36	37	36	40
32	34	34	32	33	34	31	37	33	38	36	37	38
33	32	31	33	34	33	34	36	38	36	34	35	39
34	34	32	34	33	32	33	34	35	35	38	36	37
33	33	34	34	34	34	34	37	37	38	37	34	38
31	32	31	33	32	33	32	37	35	37	35	36	39

Варіант 13.

Дані про продаж чоловічих костюмів за розмірами.

50	50	50	50	48	48	52	50	52	52	48	48
50	52	50	48	52	50	50	52	50	48	52	54
50	54	48	48	50	50	52	54	48	48	50	48
50	52	56	56	50	46	48	48	56	46	50	52
46	48	48	50	50	54	50	56	48	50	50	50
54	50	56	48	56	48	48	44	50	46	56	52
48	48	44	50	46	54	48	50	50	48	48	52
54	48	50	50	48	48	50	48	48	50	48	48
48	56	48	48	50	50	48	48	52	48	48	54
48	50	52	48	50	48	48	56	52	50	52	48

Варіант 14.

Продаж чоловічих зимових черевиків за розмірами.

40	40	38	42	40	41	42	39	38	42	41	43	42
42	42	41	42	39	38	42	42	41	44	41	41	44
44	41	40	40	39	42	44	41	40	41	42	44	41
40	39	42	44	41	40	40	39	42	42	41	42	42
42	40	43	41	42	40	42	40	43	42	40	42	42
41	36	40	42	40	43	41	42	40	43	41	43	43
41	38	39	38	43	41	36	38	39	41	40	42	41
38	43	41	41	38	39	38	43	41	40	41	41	40
43	43	41	44	42	42	43	43	41	42	40	42	42
44	42	42	43	43	41	44	42	42	42	40	41	40
40	41	40	42	44	43	40	42	40	42	43	43	43
36	44	43	42	41	36	42	44	43	41	42	41	42
42	37	42	40	41	39	42	37	42	42	43	43	
40	41	39	42	37	42	40	41	39	41	42	42	

Варіант 15.

Дані про продаж жіночих босоніжок за розмірами.

30	30	32	32	34	32	33	35	36	37	34	32	33
32	31	34	31	33	33	34	37	37	36	38	34	32
34	32	31	34	32	34	32	36	36	39	37	33	34
31	34	33	33	34	34	34	38	38	37	35	32	33
32	32	34	31	33	34	33	36	36	36	35	31	34
34	34	32	30	34	32	31	37	33	38	36	32	33
32	31	33	34	33	33	34	36	38	36	34	34	34
34	32	34	33	32	34	33	34	35	35	36	31	32
33	34	34	34	34	33	34	37	37	38	39	33	
32	31	32	33	33	31	32	37	35	37	37	34	
35	35	36	32	33	30	33	33	34	33	38	34	
36	37	40	33	32	31	34	31	33	34	39	33	
37	36	39	34	34	32	31	34	32	32	33	34	
38	38	38	34	30	34	33	33	34	34	34	33	
37	36	40	34	32	33	34	31	33	33	33	34	
30	37	38	32	34	34	32	33	34	31	31	32	

Варіант 16.

Дані про продаж дитячого взуття за розмірами.

30	34	29	33	34	32	33	33	33	34	33	33	34	33
32	31	34	31	33	33	34	32	34	33	34	34	33	34
34	32	29	34	32	34	32	34	31	32	33	34	34	33
31	34	33	30	34	34	34	31	33	34	29	33	29	31
32	33	34	31	33	34	33	32	34	33	34	33	33	32
34	34	32	33	32	34	31	34	32	34	32	34	34	32
30	31	33	34	33	33	34	29	33	29	34	33	31	34
34	32	34	29	32	34	33	34	34	32	33	34	32	30
33	34	34	34	34	33	34	33	34	34	32	32	34	32
32	31	33	32	33	31	32	32	33	33	32	31	31	30
32	33	34	30	33	34	33	32	33	32	31	33	34	34
33	32	34	34	31	33	34	34	31	33	32	32	33	31
34	34	32	31	34	32	32	31	34	34	34	34	34	30
34	31	34	33	33	34	34	32	33	34	31	33	32	33
34	32	33	34	31	33	33	34	31	34	34	34	33	32

Варіант 17.

Кількість автомобілів проданих у магазині за тиждень, штук.

30	34	30	33	29	32	33	34	34	34	33	32	33	32
32	30	34	31	33	33	34	33	33	34	34	34	32	34
34	32	31	34	30	34	32	31	30	33	32	33	31	34
31	34	30	33	34	34	34	32	33	33	33	34	33	32
32	33	34	31	33	29	29	33	32	34	31	33	34	33
34	34	32	33	34	32	31	28	34	31	34	32	33	34
32	31	33	34	33	33	34	34	31	33	33	34	34	31
34	32	30	33	32	34	28	29	32	34	31	33	34	31
33	34	34	34	34	33	34	32	34	32	33	34	33	34
32	31	33	32	33	31	32	33	32	33	34	33	34	
32	33	34	33	33	34	33	34	34	33	34	32	33	
33	32	31	29	31	33	34	33	33	34	34	34	33	
34	29	32	31	34	32	32	31	32	33	32	33	34	
34	31	34	33	33	34	34	32	31	34	33	34	32	
34	32	33	34	31	33	33	34	28	31	34	32	32	

Варіант 18.

Оцінки, які отримали студенти 4-го курсу під час сесії.

5, 4, 3, 3, 4, 4, 4, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3, 3,
2 , 5, 2, 4, 4, 4, 5, 3, 2, 2, 4, 4, 4, 5, 5, 5, 4, 3, 3, 4, 4, 4, 5, 4, 3,2, 5,
4, 4, 4, 4, 4, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3, 5, 4, 3,
3, 4, 4, 4, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3, 3, 2 , 5,
2, 4, 4, 4, 5, 3, 2, 2, 4, 4, 4, 5, 5, 5, 4, 3, 3, 4, 4, 4, 5, 4, 3,2, 5, 4, 4,
4, 4, 3, 3, 4, 4, 4, 5, 5, 2, 2, 4, 4, 4, 3, 4, 5, 5, 4, 4, 4, 5, 5, 5, 3, 3.

Варіант 19.

Число працівників цеху, що виходять щоденно на роботу.

30	34	35	33	35	32	33	34	34	32	33	34	32	31
32	31	34	31	33	33	34	32	30	33	34	35	33	34
34	32	31	34	32	34	32	34	32	34	33	32	34	33
31	34	33	33	34	34	34	33	34	35	34	34	33	34
32	35	34	31	33	35	33	32	31	33	32	33	31	32
34	34	32	30	34	32	30	32	33	34	33	33	33	34
32	31	33	34	34	33	33	33	32	31	34	31	33	34
34	32	34	33	32	34	33	30	34	32	31	34	32	32
33	35	34	34	34	33	34	34	31	34	33	33	34	34
32	31	33	32	33	30	32	34	32	30	34	31	35	33
33	34	33	33	34	32	33	32	34	34	32	33	34	31
32	31	34	31	33	33	34	33	32	31	33	34	33	34
34	35	31	34	32	34	32	34	34	32	34	33	32	33
31	34	33	30	34	34	34	33	33	34	34	34	34	34
32	35	34	31	33	34	33	31	32	31	33	32	33	32

Варіант 20.

Число поїздів, які прибувають щоденно на станцію.

32	32	32	32	31	30	33	34	30	34	32	33	32	33
33	32	31	34	31	33	34	34	34	33	34	29	34	34
34	34	32	28	34	32	32	32	32	32	33	30	32	33
34	29	29	33	33	34	34	33	33	33	31	31	33	34
34	32	33	34	31	33	33	29	34	34	34	32	34	32
32	34	34	32	28	34	31	33	34	32	34	34	31	33
33	32	28	33	34	33	34	31	33	34	33	30	29	33
34	34	32	34	33	32	33	30	33	33	34	33	32	31
33	33	34	34	34	34	34	32	31	32	33	32	33	34
31	32	31	33	32	33	32	34	33	34	32	29	31	30
33	34	33	33	34	32	33	32	34	34	33	33	33	33
32	28	34	31	33	33	34	34	33	34	34	34	34	29
34	32	31	34	32	34	32	33	34	34	33	33	34	32
31	34	33	33	28	34	34	32	32	31	34	34	33	33

Варіант 21.

Число студентів 2-го курсу, які мали пропуски занять у 1-му семестрі.

20	25	30	20	30	21	22	30	23	24	22	26	25
25	25	30	21	29	21	22	30	28	26	23	27	25
21	21	25	22	28	21	24	20	30	26	24	28	23
24	20	25	23	27	22	26	20	30	26	25	29	23
24	20	25	24	26	22	26	20	26	20	23	23	24
23	24	22	25	26	23	25	24	20	20	24	23	22

Варіант 22.

Оцінки, отримані студентами1-го курсу під час сесії.

4	5	4	3	4	4	3	4	4	5	3	3	4	3	2	4
4	5	4	2	4	4	3	4	4	4	4	3	3	5	2	3
4	4	4	2	4	5	4	4	5	4	4	2	3	3	5	3
5	4	4	5	4	3	2	5	3	5	2	3	4	4	3	3
5	4	5	5	5	3	4	5	2	3	2	3	4	4	4	4
2	4	4	5	5	2	4	5	5	2	5	4	3	3	4	4

Варіант 23.

Число дітей в одній сім’ї.

3	1	2	6	2	1	2	4	5	2	1	3	3	2	2	1
3	1	2	4	2	0	2	4	7	2	1	2	3	2	1	1
2	5	2	3	2	0	1	3	4	3	1	0	2	0	3	2
2	3	1	2	3	0	1	3	3	3	0	6	4	1	1	1
2	2	1	0	2	2	1	2	3	1	2	1	5	7	2	1
1	2	1	0	2	2	1	2	2	1	1	1	6	1	3	1

Варіант 24.

Число кімнат в одній квартирі зданого будинку.

3	5	3	1	4	2	5	1	5	4	3	2	3	2	2	2	2	2	3	4
3	5	3	1	4	2	2	1	5	4	3	3	2	3	2	2	2	3	1	2
3	1	3	1	4	5	3	1	2	2	4	3	1	3	4	3	3	4	5	1
3	1	4	2	2	1	3	2	2	2	3	4	2	3	3	3	3	5	3	5
2	1	4	2	2	1	3	2	2	2	3	4	4	3	2	1	1	4	2	3

Варіант 25.

Число спільно проживаючих членів сім'ї.

2	2	4	3	5	3	5	4	3	4	4	2	6	1	1	3	2	4	4	1	6	6	2
2	1	4	3	5	3	6	4	3	4	4	2	6	3	2	3	1	4	4	1	5	2	4
4	1	4	3	4	3	7	4	3	5	4	4	6	3	6	2	5	2	4	2	7	2	4
6	3	3	3	4	2	7	6	1	5	5	4	3	1	2	1	4	3	5	2	5	3	1
4	3	3	2	4	5	3	2	2	5	5	3	5	1	5	1	3	4	2	4			

Варіант 26.

Оцінки, отримані студентами 3-го курсу під час сесії.

5	4	3	4	5	4	3	4	4	5	2	4	3	5	4	2	4	5	5	2
4	5	4	2	4	4	3	4	4	4	3	3	4	3	5	4	4	3	4	4
4	4	4	2	4	5	4	4	5	4	3	3	4	2	3	4	3	3	4	4
5	4	4	5	4	3	2	5	3	5	4	3	3	2	3	4	5	4	3	2
5	4	5	5	5	3	4	5	2	3	4	2	3	5	4	3	3	4	4	5

Варіант 27.

Число учнів у класах шкіл району (чоловік).

30	32	34	31	32	35	30	34	30	34	31	32	36
34	30	33	31	32	35	30	34	31	34	31	32	36
36	31	33	32	32	34	30	34	32	36	30	34	32
30	31	33	33	30	34	31	34	33	32	33	32	33
32	31	34	34	34	34	31	30	34	32	33	32	32
32	32	34	35	35	32	32	30	35	30	33	32	30
32	32	33	36	36	32	33	32	34	35	35	35	

Варіант 28.

Число хлопців у класах шкіл району.

15	11	16	11	14	14	10	16	12	16
15	12	16	11	15	15	10	16	12	14
16	13	16	11	16	14	10	16	13	13
17	13	16	12	17	13	10	13	13	12
17	14	15	12	17	12	11	13	13	11
12	14	13	12	17	11	12	13	14	11
12	14	11	12	12	10	13	14	15	12
12	15	11	12	13	10	14	14	16	11
10	15	12	13	13	10	15	12	17	10
11	15	12	13	13	10	16	12	17	12

Варіант 29.
Число дівчат у класах шкіл району.

16	12	15	13	17	14	18	15	13	15
15	12	16	11	15	15	10	16	12	14
16	15	15	11	16	14	10	16	13	13
17	13	16	12	17	13	10	13	13	12
17	14	15	12	17	12	11	13	13	13
13	14	13	13	18	11	12	13	14	11
12	15	11	12	12	10	13	14	15	12
12	15	11	12	13	10	14	14	16	11
10	12	16	13	19	10	15	12	17	10
12	15	12	13	13	10	16	12	17	12

Варіант 30.

Кількість спільно проживаючих членів сім'ї (людей)

2	6	3	4	2	2	2	4	1	4	4	3	5	6	4	3	7
2	6	3	4	2	2	3	5	1	1	5	2	4	6	5	3	5
4	5	3	4	3	7	3	4	2	1	58	2	4	2	4	3	4
4	5	2	6	3	6	4	4	3	4	1	3	6	3	5	6	6
6	4	2	6	3	6	4	3	3	5	1	3	2	4	7	31	
3	4	5	7	5	5	6	3	1	5		3	1	5	4		

Приклад виконання 0-го варіанта

Дана вибірка таблицею 1.

Таблиця 1.

2	4	2	4	3	3	2	0	6	1	2	3	2	2	4	3	3	5	1	3
0	2	4	3	2	3	3	1	3	3	3	1	1	2	3	1	4	3	1	2
7	4	3	4	2	2	3	3	1	4	3	1	4	5	3	4	2	4	5	2
3	6	4	1	3	4	1	3	1	0	0	4	6	4	7	4	1	3	3	

n = 79. Перша варіанта 0. Довжина інтервала 1.

Розв'язання.

1. Складемо **варіаційний** ряд, впорядкувавши варіанти з таблиці 1 за зростанням.

Таблиця 2.

0	1	1	2	3	3	3	4	4	5
0	1	2	2	3	3	3	4	4	5
0	1	2	2	3	3	3	4	4	6
0	1	2	2	3	3	3	4	4	6
1	1	2	2	3	3	3	4	4	6
1	1	2	2	3	3	3	4	4	7
1	1	2	2	3	3	3	4	4	7
1	1	2	3	3	3	4	4	5	

2. Складемо **статистичний** ряд за допомогою таблиці 2.

У таблиці 2 варіанти x_i приймають 8 різних значень: 0, 1, 2, 3, 4, 5, 6, 7. Кожна з варіант повторюється по кілька разів, число повторень варіант – це число $n_i(i=0,7)$. Так, наприклад, варіанті $x_0 = 0$ відповідає частота $n_0 = 4$, для $x_1 = 1$ частота $n_1 = 13$ і т.д. Статистичний ряд містить значення варіант x_i і відповідні їм частоти n_i або відносні частоти n_i/n,

де $n = n_1 + n_2 + n_3 + \ldots + n_m = 4 + 13 + 14 + 24 + 16 + 3 + 3 + 2 = 79$.

Таблиця 3

Варіанти x_i	Частоти n_i	$n_i \cdot x_i$	$n_i \cdot x_i^2$	Відносні частоти n_i/n	Накопичені частоти
0	4	0	0	$\frac{4}{79}$	$\frac{4}{79}$

1	13	13	13	$\frac{13}{79}$	$\frac{17}{79}$
2	14	28	56	$\frac{14}{79}$	$\frac{31}{79}$
3	24	72	216	$\frac{24}{79}$	$\frac{55}{79}$
4	16	64	256	$\frac{16}{79}$	$\frac{71}{79}$
5	3	15	75	$\frac{3}{79}$	$\frac{74}{79}$
6	3	18	108	$\frac{3}{79}$	$\frac{77}{79}$
7	2	14	98	$\frac{2}{79}$	1
Σ	79	224	822	1	

Знаходимо середнє арифметичнепо формулі

$$\overline{x} = \frac{1}{79}\sum_{i=0}^{7} n_i \cdot x_i = \frac{224}{79} = 2.84$$

Знаходимо незміщену вибіркову дисперсію по формулі

$s_x^2 = \frac{1}{n-1}\sum_{i=0}^{7}(x_i - \overline{x})^2$,
або

$s_x^2 = \frac{n}{n-1}\left(\overline{x^2} - \overline{x}^2\right)$,
де

$\overline{x^2} = \frac{1}{79}\sum_{i=0}^{7} n_i \cdot x_i^2$

$$s_x^2 = \frac{79}{78}(10.405 - 2.84^2) = 2.396$$

Стандартне відхилення знаходимо по формулі

$$s_x = \sqrt{s_x^2} = 1.548,$$

в) медіана вибірки при непарному n – це значення **серединної варіанти** варіаційного ряду (див. табл. 3). В таблиці 3 на 39-му місці значення варіанти 3

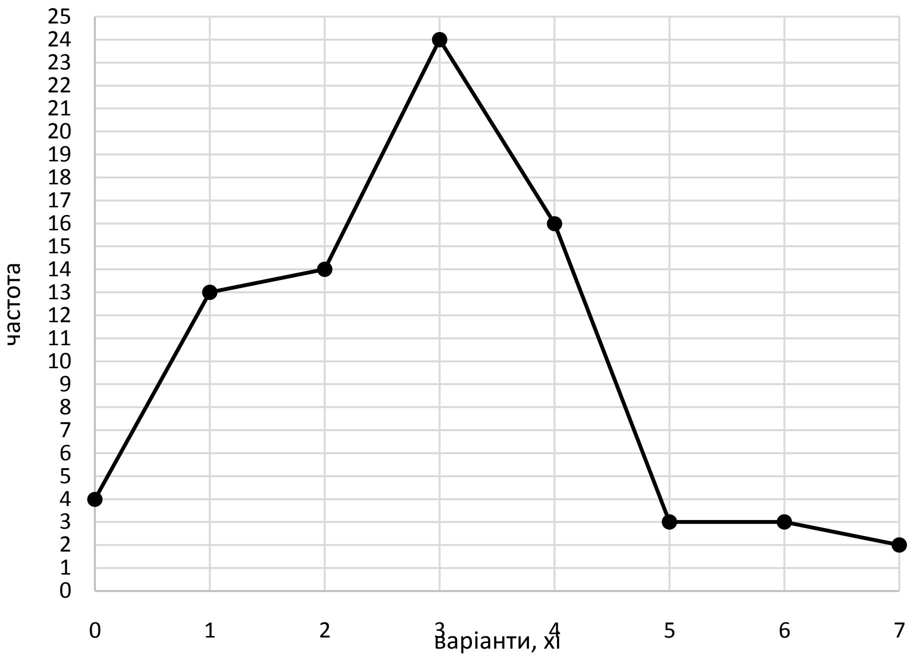
$Me = 3$,

г) Модою є значення x_i, яке має максимальну частоту, тобто

$$Mo = 3.$$

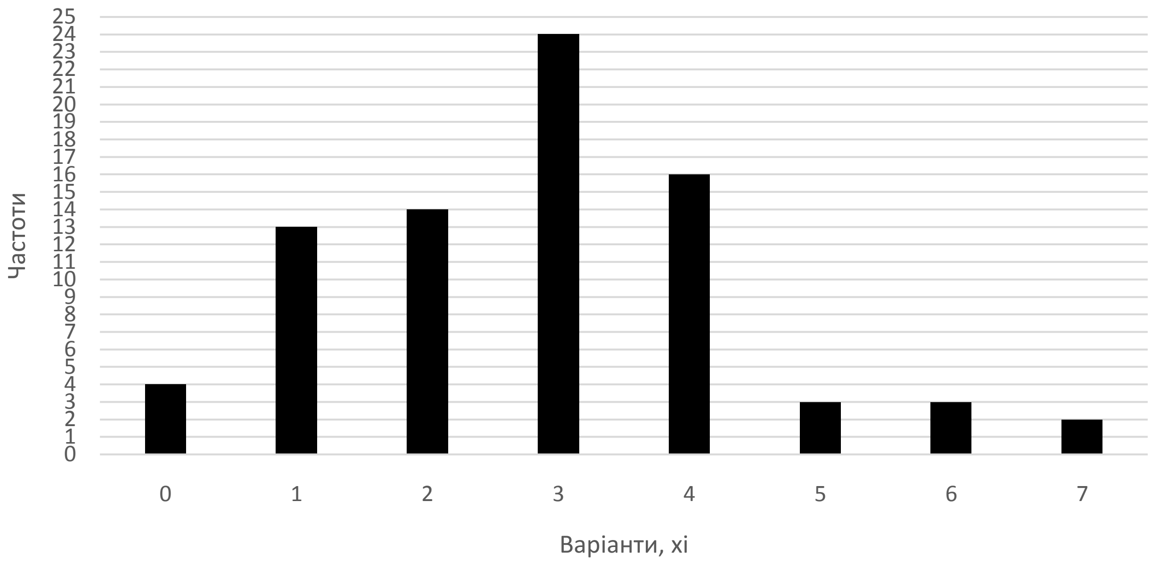
3 . а) Будуємо полігон частот.

На осі абсцис відкладаємо значення варіантx_i, а по осі ординат значення відносних частотn_i, взятих із таблиці 3.



б) Будуємо гістограму частот.

На осі абсцис відкладаємо значення варіантx_i, а по осі ординат прямокутники висотою, рівному значенню частотn_i, взятих із таблиці 3.



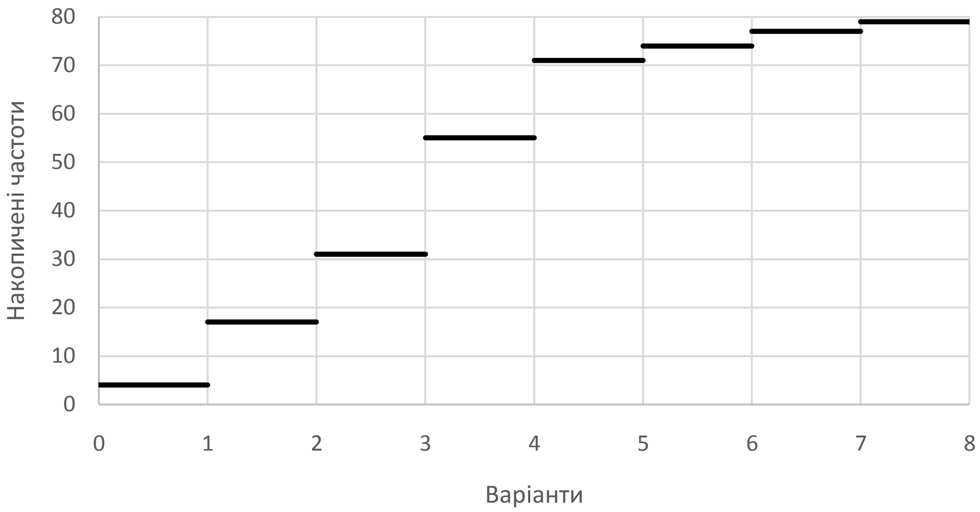
Для побудови гістограми відносних частот треба кожну ординату гістограми частот поділити на об'єм вибірки, тобто на n=79.

в) Будуємо функцію відносних накопичених відносних частот та її графік.

Функцію $F^*(x)$відносних накопиченних частот розглядаємо на тих же інтервалах, що й гістограму, тільки висота прямокутників дорівнює значенню накопиченої частоти на даному інтервалі

$$F^*(x) = \begin{cases} 0, x \leq 0 \\ \frac{4}{79}, 0 < x \ \leq 1 \\ \frac{17}{79}, 1 < x \leq 2 \\ \frac{31}{79}, 2 < x \leq 3 \\ \frac{55}{79}, 3 < x \leq 4 \\ \frac{71}{79}, 4 < x \leq 5 \\ \frac{74}{79}, 5 < x \leq 6 \\ \frac{77}{79}, 6 < x \leq 7 \\ 1, \quad 7 < x \end{cases}$$

Будуємо графік емпирічної функції накопичених частот



Для побудови графіка емпирічної функції відносних накопичених частот треба кожну ординату графіка функції накопичених частот поділити на об'єм вибірки, тобто на n=79.

4.Знайти довірчі інтервали для математичного сподівання при довірчій ймовірності γ

$$\gamma = \begin{cases} 0.8, \text{якщо } V \le 10 \\ 0.95, \text{якщо } 10 < V \le 20 \\ 0.98, \text{якщо } 20 < V \end{cases},$$

де V- номер варіанту.

Скористаємось формулою

$P\left(\overline{x} - t_\gamma \frac{s_x}{\sqrt{n}} \le m_x \le \overline{x} + t_\gamma \frac{s_x}{\sqrt{n}}\right)=\gamma$

Враховуючи, що V=0, маємо $\gamma = 0{,}8$. Значення t_γзнаходимо по таблиці розподілу Стьюдента на перетині стовпця 1-γ і рядка k=n-1. В нашому випадку $1 - \gamma = 1 - 0.8 = 0.2$, k=n-1=8-1=7, $t_{0.2}(7) = 1.41$.

Врахоуаючи, що $\overline{x} = 2.84$, $s_x = 1.548$, знаходимо

$P\left(2{,}84 - 1{,}41 \frac{1{,}548}{\sqrt{8}} \le m_x \le 2{,}84 + 1{,}41 \frac{1{,}548}{\sqrt{8}}\right)$=0.8,

або

$P(2{,}426 \le m_x \le 3{,}612)$=0.8,

Таким чином, довірчий інтервал для математичного сподівання при довірчій ймовірності 0.8 становить

$2{,}426 \le m_x \le 3{,}612.$

5.При рівні значущості α перевірити гіпотезу про нормальний закон розподілу відповідної генеральної сукупності, користуючись критерієм згоди Колмогорова, якщо

$$\alpha = \begin{cases} 0.1, \text{якщо } V \le 10 \\ 0.05, \text{якщо } 10 < V \le 20 \\ 0.02, \text{якщо } 20 < V \end{cases}$$

Зведемо необхідні значення в таблицю.Враховуємо, що

$\overline{x} = 2.84$, $s_x = 1.548$.

i	x_i	n_i	$F^*(x_i)$	$\widehat{x_i} = \frac{x_i - \overline{x}}{s_x}$	$z = \frac{1}{2}\Phi(\widehat{x_i})$	$F(x_i) = \frac{1}{2}$+ z	$\lvert F^*(x_i) - F(x_i)\rvert$
1	0	4	$\frac{4}{79}$	-1.83	- 0.4664	0.0336	0.0170
2	1	13	$\frac{17}{79}$	-1.19	- 0.3830	0.117	0.0982
3	2	14	$\frac{31}{79}$	-0.54	- 0.2054	0.2946	0.0978
4	3	24	$\frac{55}{79}$	0.10	0.0398	0.5398	**<u>0.1564</u>**
5	4	16	$\frac{71}{79}$	0.75	0.2734	0.7734	0.12853
6	5	3	$\frac{74}{79}$	1.40	0.4192	0.9192	0.0175
7	6	3	$\frac{77}{79}$	2.04	0.4793	0.9793	0.0046
8	7	2	1	2.69	0.4964	0.9964	0.0036
Σ		79					

Знаходимо

$$\lambda_0 = max|F^*(x_i) - F(x_i)| \cdot \sqrt{n} = 0.1564 \cdot \sqrt{8} = 0.442$$

За таблицею значень функції $P(\lambda)$знаходимо 1- $P(0.442)$ =0.97.

Отримана ймовірність досить висока. Тому, згідно критерія Колмогорова, розбіжність між емпирічною і теоретичною функціями розподілу не є істотною, її можна пояснити випадковостю. Таким чином, можна стверджувати, що представлені дані розподілені за нормальним законом.

Приклади тестових завдань

Середнє значення X обчислюється за формулою

A. $\sum_{i=1}^{n} x_i$ B. $\frac{1}{n}\sum_{i=1}^{n} x_i$ C. $\frac{1}{n-1}\sum_{i=1}^{n} x_i$ D. інша відповідь

Незміщена вибіркова дисперсія обчислюється за формулою

A. $\frac{1}{n}\sum_{i=1}^{n} x_i^2$ B. $\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$ C. $\frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2$ D. інша відповідь

Коефіцієнт парної кореляції r_{yx} обчислюється за формулою

A. $\overline{xy} - \bar{x} \cdot \bar{y}$ B. $\frac{\overline{xy} - \bar{x} \cdot \bar{y}}{s_x s_y}$ C. $\frac{\overline{xy} - \bar{x} \cdot \bar{y}}{n \cdot s_x s_y}$ D. інша відповідь

Вибрати неправильне твердження – зміст МНК полягає в

A. мінімізації суми квадратів параметрів регресії,

B. мінімізації суми квадратів значень залежної змінної,

C. мінімізації суми квадратів відхилень фактичних даних від обчислених з рівняння регресії

D. всі варіанти правильні.

Нульова гіпотеза позначається

A. H_0 B. H_1 C. H_2 D. інша відповідь

Додатки

Таблиця 1

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,71	0,2611	1,41	0,4207	2,22	0.4868
0,01	0,0040	0,72	0,2642	1,42	0,4222	2,24	0,4875
0,02	0,0080	0,73	0,2373	1,43	0,4236	2,26	0,4881
0,03	0,0120	0,74	0,2703	1,44	0,4251	2,28	0,4887
0,04	0,0160	0,75	0,2734	1,45	0,4265	2,30	0,4893
0,05	0,0199	0,76	0,2764	1,46	0,4279	2,32	0,4898
0,06	0,0239	0,77	0,2794	1,47	0,4292	2,34	0,4904
0,07	0,0279	0,78	0,2823	1,48	0,4306	2,36	0,4909
0,08	0,0319	0,79	0,2852	1,49	0,4319	2,38	0,4913
0,09	0,0359	0,80	0,2881	1,50	0,4332	2,40	0,4918
0,10	0,0398	0,81	0,2910	1,51	0,4345	2,42	0,4922
0,11	0,0438	0,82	0,2939	1,52	0,4357	2,44	0,4927
0,12	0,0478	0,83	0,2967	1,53	0,4370	2,46	0,4931
0,13	0,0517	0,84	0,2995	1,54	0,4382	2,48	0,4934
0,14	0,0557	0,85	0,3023	1,55	0,4394	2,50	0,4938
0,15	0,0596	0,86	0,3051	1,56	0,4406	2,52	0,4941
0,16	0,0636	0,87	0,3078	1,57	0,4418	2,54	0,4945
0,17	0,0675	0,88	0,3106	1,58	0,4429	2,56	0,4948
0,18	0,0714	0,89	0,3133	1,59	0,4441	2,58	0,4951
0,19	0,0753	0,90	0,3159	1,60	0,4452	2,60	0,4953
0,20	0,0793	0,91	0,3186	1,61	0,4463	2,62	0,4956
0,21	0,0832	0,92	0,3212	1,62	0,4474	2,64	0,4959
0,22	0,0871	0,93	0,3238	1,63	0,4484	2,66	0,4961
0,23	0,0910	0,94	0,3264	1,64	0,4495	2,68	0,4963
0,24	0,0948	0,95	0,3289	1,65	0,4505	2,70	0,4965
0,25	0,0987	0,96	0,3315	1,66	0,4515	2,72	0,4967
0,26	0,1026	0,97	0,3340	1,67	0,4525	2,74	0,4969
0,27	0,1064	0,98	0,3365	1,68	0,4535	2,76	0,4971
0,28	0,1103	0,99	0,3389	1,69	0,4545	2,78	0.4973
0,29	0,1141	1,00	0,3413	1,70	0.4554	2,80	0,4974
0,30	0,1179	1,01	0,3438	1.71	0,4564	2,82	0,4976
0,31	0,1217	1,02	0,3461	1,72	0.4573	2,84	0,4977
0,32	0,1255	1,03	0,3485	1,73	0,4582	2,86	0,4979
0,33	0,1293	1,04	0,3508	1,74	0,4591	2,88	0,4980
0,34	0,1331	1,05	0,3531	1,75	0.4599	2,90	0,4981
0,35	0,1368	1,06	0,3554	1,76	0.4608	2,92	0,4982
0,36	0,1406	1,07	0,3577	1,77	0,4616	2,94	0,4984
0,37	0,1443	1,08	0,3599	1,78	0.4625	2,96	0,4985
0,38	0,1480	1,09	0,3621	1,79	0,4633	2,98	0,4986
0,39	0,1517	1,10	0,3643	1,80	0,4641	3,00	0,49865

0,40	0.1554	1,11	0,3665	1,81	0,4649	3,20	0,49931
0,41	0,1591	1,12	0,3686	1,82	0.4656	3,40	0,49966
0,12	0,1628	1,13	0,3708	1,83	0,4664	3,60	0,499841
0.43	0,1664	1,14	0,3729	1,84	0.4671	3,80	0,499928
0,44	0,1700	1,15	0,3749	1,85	0,4678	4,00	0,499968
0,45	0,1736	1,16	0,3770	1,86	0,4686	4,50	0,499997
0,46	0,1772	1,17	0,3790	1,87	0,4693	5,00	0,499997
0,47	0,1808	1,18	0,3810	1,88	0,4699		
0,48	0,1844	1,19	0,3830	1,89	0,4706		
0,49	0,1879	1,20	0,3849	1,90	0,4713		
0,50	0,1915	1,21	0,3869	1,91	0,4719		
0,51	0,1950	1,22	0,3883	1,92	0,4726		
0,52	0,1985	1,23	0,3907	1,93	0,4732		
0,53	0,2019	1,24	0,3925	1,94	0,4738		
0,54	0,2054	1,25	0,3944	1,95	0,4744		
0,55	0,2088	1,26	0,3962	1,96	0,4750		
0,56	0,2123	1,27	0,3980	1,97	0,4756		
0,57	0,2157	1,28	0,3997	1,98	0,4761		
0,58	0,2190	1.29	0,4015	1,99	0,4767		
0,59	0,2224	1.30	0,4032	2,00	0,4772		
0,60	0,2257	1,31	0,4049	2,02	0,4783		
0,61	0,2291	1,32	0,4066	2,04	0,4793		
0,62	0,2324	1,33	0,4082	2,06	0,4803		
0,63	0,2357	1,34	0,4099	2,08	0,4812		
0,64	0,2389	1,35	0,4115	2,10	0,4821		
0,65	0,2422	1,36	0,4131	2,12	0,4830		
0,66	0,2454	1,37	0,4147	2,14	0,4838		
0,67	0,2486	1,38	0,4162	2,16	0,4846		
0,68	0,2517	1,39	0,4177	2,18	0,4854		
0,69	0,2549	1,40	0,4192	2,20	0,4861		
0,70	0,2580						

Таблиця 2

Квантилі розподілу Стьюдента $t_{\alpha}(n)$

α / n	0,80	0,60	0,50	0,40	0,20	0,10	0,05	0,02	0,01
1	2	3	4	5	6	7	8	9	10
1	0,325	0,727	1,000	1,376	3,078	6,314	12,706	31,821	63,657
2	0,289	0,617	0,816	1,061	1,886	2,920	4,303	6,965	9,925
3	0,277	0,584	0,765	0,978	1,638	2,353	3,182	4,541	5,841
4	0,271	0,569	0,741	0,941	1,533	2,132	2,776	3,747	4,604
5	0,267	0,559	0,727	0,920	1,476	2,015	2,571	3,365	4,032
6	0,265	0,553	0,718	0,906	1,440	1,943	2,447	3,143	3,707
7	0,263	0,549	0,711	0,896	1,415	1,895	2,365	2,998	3,499
8	0,262	0,546	0,706	0,889	1,397	1,860	2,306	2,896	3,355
9	0,261	0,543	0,703	0,883	1,383	1,833	2,262	2,821	3,250
10	0,260	0,542	0,700	0,879	1,372	1,812	2,228	2,764	3,169
11	0,260	0,540	0,697	0,876	1,363	1,796	2,201	2,718	3,306
12	0,259	0,539	0,695	0,873	1,356	1,782	2,179	2,681	3,055
13	0,259	0,538	0,694	0,870	1,350	1,771	2,160	2,650	3,012
14	0,258	0,537	0,692	0,868	1,345	1,761	2,145	2,624	2,977
15	0,258	0,536	0,691	0,866	1,341	1,753	2,131	2,602	2,947
16	0,258	0,535	0,690	0,865	1,337	1,746	2,120	2,583	2,921
17	0,257	0,534	0,689	0,863	1,333	1,740	2,110	2,567	2,898
18	0,257	0,534	0,688	0,862	1,330	1,734	2,101	2,552	2,878
19	0,257	0,533	0,688	0,861	1,328	1,729	2,093	2,539	2,861
20	0,257	0,533	0,687	0,860	1,325	1,725	2,086	2,528	2,845
21	0,257	0,532	0,686	0,859	1,323	1,721	2,080	2,518	2,831
22	0,256	0,532	0,686	0,858	1,321	1,717	2,074	2,508	2,819
23	0,256	0,532	0,685	0,858	1,319	1,714	2,069	2,500	2,807
24	0,256	0,531	0,685	0,857	1,318	1,711	2,064	2,492	2,797
25	0,256	0,531	0,684	0,856	1,316	1,708	2,060	2,485	2,787
26	0,256	0,531	0,684	0,856	1,315	1,706	2,056	2,479	2,779
27	0,256	0,531	0,684	0,855	1,314	1,703	2,052	2,473	2,771
28	0,256	0,530	0,683	0,855	1,313	1,701	2,048	2,467	2,763

29	0,256	0,530	0,683	0,854	1,311	1,699	2,045	2,462	2,756
1	2	3	4	5	6	7	8	9	10
30	0,256	0,530	0,683	0,854	1,310	1,697	2,042	2,457	2,750
40	0,255	0,529	0,681	0,851	1,303	1,684	2,021	2,423	2,704
60	0,254	0,527	0,679	0,848	1,296	1,671	2,000	2,390	2,660
100	0,254	0,526	0,677	0,845	1,290	1,660	1,984	2,364	2,626
200	0,254	0,525	0,676	0,843	1,286	1,652	1,972	2,345	2,601
∞	0,253	0,524	0,675	0,842	1,282	1,645	1,960	2,326	2,576

Література

1. Андрухаев Х.М. Сборник задач по теории вероятностей / Х.М.Андрухаев. – М.: «Просвещение»,1985.–165с.
2. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е.Гмурман – М.: Высшая школа, 2003.– 479с.
3. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике / В.Е. Гмурман.–М. : Высшая школа,1999.– 400с.
4. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов / Н.Ш. Кремер.– М.: ЮНИТИ-ДАНА,2003.-543 с.
5. Красс М.С. Основы математика и ее приложения в экономическом образовании / М.С. Красс, Б.П.Чупрынов.–М.: Дело, 2000.– 688с.
6. Рудавський Ю.К.,Збірник задач з теорії ймовірностей. Навч. посібник / Ю.К.Рудавський,П.П.Костробій,І.Я.Олексів.–Львів: «Львівська політехніка», 2000. – 242с.

Навчальне видання

***Серебреников** Вадим Михайлович*
***Квітка** Тетяна Володимирівна*

Кафедра загальноінженерних дисциплін та обладнання

МАТЕМАТИКА ДЛЯ ЕКОНОМІСТІВ: ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Навчально-методичний посібник

розділ «Математична статистика»

Формат 60х84/8. Ум. др. арк.

Донецький національний університет
економіки і торгівлі імені
Михайла Туган-Барановського
50042, Дніпропетровська обл.,
м.Кривий Ріг, вул. Курчатова, 13.
Свідоцтво суб’єкту видавничої
Справи ДК № 4929 від 07.07.2015 р.