190623

Міністерство освіти і науки України
Херсонський державний педагогічний університет
Кафедра математики

Кузьмич Л.В.

МАТЕМАТИЧНА СТАТИСТИКА

Методичні вказівки до вивчення курсу

(для студентів спеціальностей:
7.070801, 7.010103 Екологія. ПМСО. Географія,
7.040101, 7.010103 Психологія. ПМСО. Біологія,
7.010106 Дефектологія. Олігофренопедагогіка і логопедія,
7.010103 ПМСО. Географія і біологія.)

Херсон - 2000

Міністерство освіти і науки України
Херсонський державний педагогічний університет
Кафедра математики

Кузьмич Л.В.

МАТЕМАТИЧНА СТАТИСТИКА

Методичні вказівки до вивчення курсу

для студентів спеціальностей:
7.070801, 7.010103 Екологія. ПМСО. Географія,
7.040101 7.010103 Психологія. ПМСО. Біологія,
7.010106 Дефектологія. Олігофренопедагогіка і логопедія,
7.010103 ПМСО. Географія і біологія

Херсон – 2000

Методичні вказівки обговорено на засіданні кафедри математики (протокол №11 від "06" 06.2000 р.)

Схвалено навчально-методичною радою університету (протокол № 3 від 26.06.2000 р.)

Рекомендовано до видання Вченою радою Херсонського державного педагогічного університету (протокол № 7 від "3" 07. 2000 р.)

Укладач: Кузьмич Людмила Василівна, кандидат педагогічних наук.

Рецензенти: Крекнін В.А., кандидат фізико-математичних наук, доцент,

Самойленко В.Г., кандидат фізико-математичних наук, доцент.

Кузьмич Л.В.
Математична статистика: Методичні вказівки. –Херсон: Айлант, 2000. –36с.

ISBN 966-7403-50-5

© Кузьмич Л.В.,2000
© ХДПУ, 2000

ЗМІСТ

ПРОГРАМА КУРСУ

Ця програма визначає мінімальний обсяг знань з математичної статистики, які необхідні для якісної підготовки вчителів географії загальноосвітньої школи, працівників екологічно-географічної галузі та психологічної служби. В даних методичних вказівках містяться початкові відомості з математичної статистики. Поряд з теоретичним матеріалом пропонуються зразки розв'язування задач, питання для самоперевірки, тексти контрольних робіт, додатки з довідковими таблицями.

Основні поняття та методи, передбачені програмою, ілюструються прикладами, які пов'язані з навчальною спеціальністю студентів.

Статистична сукупність. Ряд розподілу (варіаційний ряд), його геометрична ілюстрація (полігон, гістограма). Середнє значення ознаки і середнє квадратичне відхилення, методи їх розрахунку.

Оцінка параметрів генеральної сукупності за її вибіркою. Означення статистичної оцінки. Точкові статистичні оцінки: зсунені та незсунені, ефективні, спроможні. Точність і надійність оцінки.

Дискретні та неперервні випадкові величини, закони їх розподілу.

Інтервали надійності для параметрів нормального розподілу. Означення надійного інтервалу. Побудова надійного інтервалу для математичного сподівання та середнього квадратичного відхилення. Оцінка точності вимірювань.

Перевірка статистичних гіпотез. Нульова та альтернативна, проста та складна гіпотези. Статистичний критерій. Спостережуване значення критерію. Критична область. Область прийняття нульової гіпотези; критична точка. Емпіричні та теоретичні частоти. Критерій згоди.

Елементи теорії кореляції. Лінійна кореляція. Кореляційна залежність. Коефіцієнт кореляції. Розрахунок прямих регресії.

ЛІТЕРАТУРА

1. Кудрявцев В.А., Демидович В.П. Краткий курс высшей математики. – М.: Наука, 1985.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1975.
3. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. – М.: Высшая школа, 1975.
4. Данко П.Е., Попов А.Г., Кожевникова Т.Я. Высшая математика в упражнениях и задачах. Ч.II. – М.: Высшая школа, 1986. – 416 с.
5. Солодовников А.С. Теория вероятностей. – М.: Просвещение, 1978.

6. Виленкин Н.Я. Индукция. Комбинаторика. – М.:Просвещение, 1976.
7. Баврин И.И. Высшая математика. – М.: Просвещение, 1980. – 384 с.
8. Гнеденко Б.В.Курс теории вероятностей. – М.: Наука, 1969.
9. Бейли Н. Математика в биологии и медицине. – М.: Мир, 1970.
10. Дюженкова Л.І., Носаль Т.В. Вища математика. Практикум. – К.: Вища школа, 1991. – 408 с.
11. Самнер Г. Математика для географов. Пер. с англ. – М.: Прогресс, 1981. – 296 с.
12. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматгиз, 1963.
13. Архипов ЮР, Блажко Н.И., Григорьев С.В. и др. Математические методы в географии. – Казань, 1976.

ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

1. ***Математична статистика*** – це розділ математики, в якому розробляються і вивчаються науково обґрунтовані методи збору та обробки статистичних даних.

2. ***Статистичною сукупністю*** називається множина об'єктів (однорідних), які необхідно вивчити.

3. У об'єктів будуть вивчатися як кількісні, так і якісні ознаки.

4. Для вивчення тієї чи іншої статистичної сукупності об'єктів краще зробити її повне дослідження, тобто вивчити кожний об'єкт. Але з тих чи інших причин в більшості випадків зробити повне дослідження неможливо.

5. Основними причинами, які не дозволяють зробити повне дослідження, є:

а) велика кількість об'єктів дослідження;

б) недоступність більшої частини об'єктів дослідження.

6. Якщо повне дослідження неможливе, то із всієї статистичної сукупності об'єктів для вивчення вибирають тільки її частину.

7. Статистична сукупність, з якої вибирають частину об'єктів, називається ***генеральною сукупністю***.

8. ***Вибіркою*** називається множина об'єктів, які випадково вибрані з генеральної сукупності.

9. *Об'ємом генеральної сукупності* називається число її об'єктів. *Об'ємом вибірки* називається кількість об'єктів вибірки.

10. Приклад: 200 яблук даного дерева досліджуються на наявність специфічного для даного сорту смаку. Для цього вибирають 10 яблук. Тут: об'єм генеральної сукупності – 200, об'єм вибірки – 10.

11. Якщо вибірку вибирають по одному об'єкту, який досліджується і повертається назад в генеральну сукупність, то таку вибірку називають *вибіркою з поверненням* або *вибіркою з повторенням*. Вибірка називається *вибіркою без повторень*, якщо об'єкт вибірки не повертається назад в генеральну сукупність. На практиці найчастіше використовують безповторну вибірку.

12. Якщо об'єм вибірки складає незначну частину генеральної сукупності, то повторна вибірка мало чим відрізняється від безповторної.

13. Вибірка повинна бути репрезентативною, тобто властивості об'єктів вибірки повинні правильно характеризувати властивості об'єктів генеральної сукупності. Вибірка вважається *репрезентативною*, якщо всі об'єкти генеральної сукупності мають однакову ймовірність потрапити в вибірку.

СТАТИСТИЧНИЙ РОЗПОДІЛ ВИБІРКИ. ПОЛІГОН. ГІСТОГРАМА

1. Нехай з деякої генеральної сукупності зроблена вибірка, причому:

значення x_1 спостерігається n_1 раз,
значення x_2 спостерігається n_2 раз,
значення x_3 спостерігається n_3 раз,
....
значення x_k спостерігається n_k раз,
при цьому $n_1+n_2+n_3+....+n_k=n$ – це об'єм вибірки.

2. Значення, що спостерігаються, тобто $x_1, x_2, ... , x_k$ – називаються *варіантами*.

3. Послідовність варіант, записаних у зростаючому порядку, називається *варіаційним рядом*.

4. Число спостережень відповідних значень, тобто числа n_1, n_2, n_3, ..., n_k, називається *частотою*.

5. Відношення частоти до об'єму вибірки називається *відносною частотою*.

6. Найчастіше відносна частота назначається P^* (або ω), тоді: $P_1^*=n_1/n$; $P_2^*=n_2/n$; .... ; $P_k^*=n_k/n$. Сума відносних частот завжди дорівнює одиниці: $P_1^*+P_2^*+P_3^*+.....+P_k^*=1$.

7. ***Статистичний розподіл вибірки*** – це перелік варіант і відповідних їм частот або варіант і відповідних їм відносних частот.

8. Статистичний розподіл можна задавати також і у вигляді переліку послідовних інтервалів і відповідних їм частот або відносних їм частот.

9. За частоту, що відповідає інтервалу, приймається сума частот варіант, які попали в даний інтервал.

10. Найчастіше статистичний розподіл задається у вигляді таблиці:

Варіанта	x_1	x_2	...	x_i
частота	n_1	n_2	...	n_i
відносна частота	w_1	w_2	...	w_i

11. Приклад. Перейти від частот до відносних частот у такому статистичному розподілі вибірки об'ємом $n = 20$.

Варіанта	2	6	12
частота	3	10	7

Відповідь:

варіанта	2	6	12
Відносна частота	0,15	0,50	0,35

12. Розмах варіації $R=x_{max}-x_{min}$.

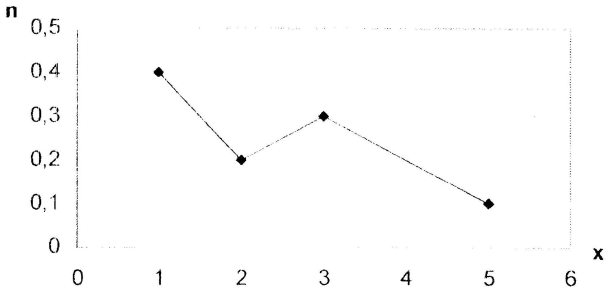
13. Для графічного зображення статистичного розподілу використовують полігон і гістограму.

14. ***Полігон***. Для побудови полігона на осі OX відкладають значення варіант x_i, на осі OY – значення частот (відносних частот) n_i (w_i).

Зауваження. При побудові полігону можуть використовуватися рівномірні і нерівномірні шкали.

15. **Приклад**. Побудувати полігон статистичного розподілу.

Варіанта	1	2	3	5
відносна частота	0,4	0,2	0,3	0,1



Цей графік – **полігон**

16. Полігоном користуються у випадку невеликої кількості варіант або у випадку дискретного розподілу ознаки вибірки.

17. У випадку великої кількості варіант або у випадку неперервного розподілу деякої ознаки користуються гістограмою.

Щоб побудувати ***гістограму***, потрібно:

а) інтервал, в якому розміщенні всі значення ознаки, що спостерігається, розбивається на декілька окремих інтервалів однакової довжини, нехай довжина їх буде h;

б) для кожного окремого інтервалу знаходять суму частот варіант, що попали в цей інтервал;

в) потім на цих інтервалах, як на сторонах будуть прямокутники з висотою n_i/h, або з висотою $n_i/(n\ h)$, де n – об'єм вибірки. Гістограма побудована.

18. Площа гістограми дорівнює сумі всіх частот (об'єму вибірки) або сумі всіх відносних частот (1).

19. **Приклад**: Побудувати гістограму неперервного розподілу об'ємом n=100.

Окремий інтервал	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Сума частот варіант	4	6	16	36	24	10	4
Висота прямокутника	0,8	1,2	3,2	7,2	4,8	2	0,8

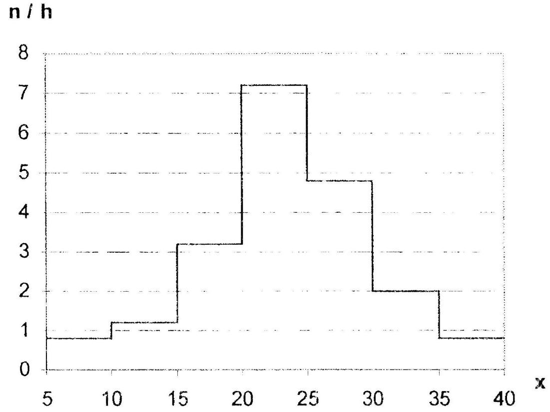
1) заданий інтервал, в якому розташовані всі варіанти, від 5 до 40.

2) інтервал варіант уже розбитий на окремі інтервали. Довжина кожного інтервалу h=5.

3) сума частот варіант кожного з окремих інтервалів вже задана.

4) обчислимо n_i/h: $n_1/h=4/5=0.8$; $n_2/h=6/5=1,2$; $n_3/h=16/5=3,2$; $n_4/h=36/5=7,2$; $n_5/h=24/5=4,8$; $n_6/h=10/5=2$; $n_7/h=4/5=0,8$.

5) Запишемо нові дані у таблицю статистичного розподілу даної вибірки (третій рядок) і побудуємо гістограму.



ОЦІНКА ПАРАМЕТРІВ ГЕНЕРАЛЬНОЇ СУКУПНОСТІ ЗА ЇЇ ВИБІРКОЮ

1. Нехай задана генеральна сукупність, кожен об'єкт якої має одну й ту саму кількісну ознаку X. При випадковому виборі об'єкта генеральної сукупності стає відомим конкретне значення кількісної ознаки X. Це значення кількісної ознаки X позначимо через *x*. При цьому X – це випадкова величина, *x* – це одне з можливих значень цієї випадкової величини.

2. Оскільки досліджується не вся генеральна сукупність, а тільки вибірка, то в розпорядженні дослідника завжди будуть тільки дані про вибірку. Тобто він матиме тільки значення варіант $x_1, x_2, x_3, \ldots, x_n$, якщо об'єм вибірки n. Через ці дані і виражаються параметри, які необхідно оцінити.

3. Найчастіше параметрами, які необхідно оцінити є:

а) математичне сподівання (МС);

в) дисперсія (Д);

г) середнє квадратичне відхилення (σ).

4. Дослідні значення $x_1, x_2, x_3, \ldots, x_n$ деякої кількісної ознаки X можна розглянути не як окремі значення, а як значення різних випадкових величин $X_1, X_2, \ldots, X_n$, з тим же розподілом, що X, тобто $M(X) = M(x)$, $Д(X) = Д(x)$.

5. Величини $X_1, X_2, \dots, X_n$ можна вважати незалежними, оскільки спостереження, які проводяться, незалежні. В цьому випадку значення $x_1, x_2, x_3, \dots, x_n$ називаються реалізаціями випадкових величин.

6. Знайти оцінку (***статистичну оцінку***) невідомого параметра теоретичного розподілу – це значить знайти функцію від випадкових величин $X_1, X_2, \dots, X_n$, що спостерігаються. Ця функція і дає наближене значення параметра, який оцінюється.

7. Статистична оцінка невідомого параметра генеральної сукупності одним числом називається ***точковою***.

8. Точкові оцінки бувають зсунені і незсунені, ефективні та спроможні (слушні).

9. Нехай Θ^* є статистична оцінка невідомого параметра Θ теоретичного розподілу. ***Зсуненою*** називають статистичну оцінку Θ^*, математичне сподівання якої не дорівнює параметру, який оцінюється. Використання такої оцінки призводить до систематичних помилок (похибок).

10. ***Незсуненою*** називають статистичну оцінку Θ^*, математичне сподівання якої дорівнює параметру, який оцінюється, тобто $M(\Theta^*)=\Theta$. Дотримання цієї вимоги усуває систематичні похибки.

11. Але незсунена оцінка не завжди дає добре наближення параметра, який оцінюється. Тому до статистичної оцінки ставиться вимога ефективності. ***Ефективною*** називають статистичну оцінку, яка (при заданому об'ємі вибірки n) має найменшу можливу дисперсію.

12. При розглядувані вибірок великого об'єму до статистичних оцінок ставиться вимога спроможності. ***Спроможною*** називають статистину оцінку, яка при $n\to\infty$ прямує по ймовірності до параметра, який оцінюється. Наприклад, якщо дисперсія незсуненої оцінки при $n\to\infty$ прямує до нуля, то така оцінка називається також спроможною.

Питання для самоперевірки

1. Чим займається математична статистика?
2. Назвіть основні задачі математичної статистики.
3. Назвіть основні поняття математичної статистики.
4. Що таке гістограма; полігон?
5. Чим відрізняються генеральна та вибіркова сукупності?
6. Назвіть параметри, які найчастіше використовують для оцінки генеральної сукупності.
7. Яка оцінка параметра називається спроможною?
8. В чому заключається різниця між генеральною сукупністю та вибіркою?
9. Яка оцінка параметра називається зсуненою? незсуненою?
10. Яка оцінка параметра називається точковою?

ЧИСЛОВІ ХАРАКТЕРИСТИКИ ГЕНЕРАЛЬНОЇ СУКУПНОСТІ ТА ВИБІРКИ. МЕТОДИ ЇХ РОЗРАХУНКУ

1. Нехай вивчається деяка генеральна сукупність об'ємом N відносно деякої кількісної ознаки X.

2. ***Генеральною середньою*** називається середнє арифметичне значення ознаки генеральної сукупності, що вивчається. Позначення генеральної середньої: **a** або $\bar{x}$.

3. Якщо всі значення кількісної ознаки, яка вивчається, різні, то генеральна середня $\bar{x} = \dfrac{1}{N}(x_1 + x_2 + \dots + x_i) = \dfrac{\sum_{i=1}^{n} x_i}{N}$.

Якщо x_1 спостерігається N_1 раз, x_2 спостерігається N_2 раз, . . ., x_n спостерігається N_i раз, то $\bar{x} = \dfrac{1}{N}(x_1N_1 + x_2N_2 + \dots + x_iN_i)$. Цю формулу можна записати компактніше: $\bar{x}$ =1/N($\sum_{i=1}^{n} x_iN_i$).

4. M(x) = $\bar{x}_г$. Цей результат справедливий, коли n_1, n_2, ... n_i – однакові і коли вони різні. У випадку неперервного розподілу деякої кількісної ознаки X за означенням вважається, що M(x) = $\bar{x}_г$.

5. Нехай для вивчення генеральної сукупності відносно деякої кількісної ознаки X зроблена вибірка об'ємом *n*. *Вибірковою середньою* називається середнє арифметичне значення ознаки вибірки даної генеральної сукупності. Позначення: $\bar{x}_в$.

6. Якщо значення x_1, x_2, x_3, ... , x_n кількісної ознаки X вибірки об'ємом *n* різні між собою, то середня вибіркова $\bar{x}_в = \dfrac{1}{n}(x_1 + x_2 + \dots + x_i)$.

7. Якщо значення $x_1, x_2, x_3, \dots, x_n$ кількісної ознаки X вибірки об'ємом *n* мають відповідні частоти n_1, n_2, ... n_i, тоді $\bar{x}_в = \dfrac{1}{n}(x_1n_1 + x_2n_2 + \dots + x_in_i)$.

8. Вибіркова середня $\bar{x}_в$ для різних вибірок одного й того ж об'єму *n* однієї і тієї ж самої генеральної сукупності в загальному випадку може бути різною.

9. Позначимо через $\bar{x}$ вибіркову середню випадкову величину, причому $\bar{x} = \frac{1}{n}(x_1 + x_2 + ... + x_i)$.

10. Обчислимо математичне сподівання вибіркової середньої:

M($\bar{x}$)=M($\frac{1}{n}(x_1 + x_2 + ... + x_i)$)=**a**.

Використовуючи властивості математичного сподівання, можна записати: a=(1/n)·(M(x_1)+ M(x_2)+ ... + M(x_i)). Використовуючи M(x_s) = M(x), запишемо: (1/n)n M (x) = M(X). Маємо: M (X) = $\bar{x}_в$ = a

Висновок: Математичне сподівання вибіркової середньої співпадає з середньою генеральною. На практиці середня арифметична ототожнюється з математичним сподіванням *a* (істинним значенням параметра, величини).

11. Якщо значення варіант x_i дуже велике число, то для полегшення обчислення вибіркової середньої використовують формулу $\bar{x} = C + \bar{u}$, де $u_i = x_i - C$, і C – константа, яка має спеціальну назву «*хибний нуль*». Значення "хибного нуля" вибирається самостійно в залежності від конкретної задачі. По можливості С беруть круглим числом (10-20 і т.д.), але таким, щоб різниці x_i – С були не великі. Найчастіше за "хибний нуль" беруть або "середнє" із значень варіант, або близьке значення до середнього арифметичного варіант, або значення варіанти, у якої найбільша частота.

12. Властивості математичного сподівання.

– Математичне сподівання постійної величини дорівнює самій постійній: M(C)=C.

– Постійний множник можна виносити за знак математичного сподівання:

M(CX)=CM(X).

– Математичне сподівання добутку взаємно незалежних випадкових величин дорівнює добутку математичних сподівань множників:

$M(X_1X_2...X_n)=M(X_1)M(X_2)...M(X_n)$.

– Математичне сподівання суми випадкових величин дорівнює сумі математичних сподівань доданків:

$M(X_1 + X_2 +...+X_n)=M(X_1) + M(X_2) +...+ M(X_n)$.

–Якщо кожну ознаку змінити на певну величину, то математичне сподівання зміниться на таку ж величину (див. п.13).

– Якщо кожну ознаку змінити у *k* разів (тобто ввести умовну варіанту *u=kx*), то математичне сподівання зміниться теж у *k* разів. Тому, знайшовши математичне сподівання умовних варіант, треба розділити його на *k*:

M(X)=(1/*k*) M(U).

13. **Приклад 1**: Вибірковим шляхом були отримані такі дані про масу 20 морських свинок при народженні (г): 30, 30, 25, 32, 30, 25, 33, 32, 29, 28, 27, 36, 31, 34, 30, 23, 28, 31, 36, 30. Знайти вибіркову середню.

Розв'язування: $\bar{x} = \frac{1}{n}(x_1 + x_2 + ... + x_i)$ = (30 ·5 +25· 2 +32· 2 +33 +29 +28· 2 +27 +36 · ·2+31· 2 +34 +23) / 20=600 / 20=30.

Приклад 2: Дана вибірка: x_1 =71,88; x_2 = 71,93; x_3 = 72,05; x_4 = 72,07; x_5=71,90; x_6 =72,02; x_7= =71,93 ; x_8=71,77; x_9 = 72,71; x_{10} =71,96. Знайти вибіркову середню.

Розв'язування Щоб не оперувати великими числами, використовуємо формулу: $\bar{x}$ =C+$\bar{u}$, де $u_i=x_i$ – **С**. В ролі хибного нуля виберемо С=72,0.

Визначимо різниці виду x_i-С.

u_1= -0,12; u_2= -0,07; u_3=0,05; u_4=0,07; u_5= -0,1; u_6=0,02; u_7 = -0,07; u_8= -0,23, u_9 =0,71; u_{10}= -0,04.

Знайдемо: $\bar{x}_в = \frac{u_1n_1 + u_2n_2 + ... + u_in_i}{n} + C = \bar{u} + C$.

$$\bar{x}_в = \frac{-0.12 - 0.07 + 0.05 + 0.07 - 0.1 + 0.02 - 0.01 - 0.23 + 0.71 - 0.04 = 0.27}{10} + 72.0 \approx$$

$\approx 0.2 + 72.0 = 72.2$

14. Для характеристики розсіювання спостережуваних значень кількісної ознаки вибірки навколо свого середнього значення ($\bar{x}_в$) застосовують таку характеристику, як вибіркова дисперсія.

Вибірковою дисперсією ($D_в$) називається середнє арифметичне квадрату відхилення спостережуваних значень ознаки розподілу X від вибіркової середньої $\bar{x}_в$:

$D_в = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_в)^2$, якщо всі x_i різні.

Якщо варіанти x_i появляються відповідно n_1, n_2,..., n_i разів, де n_1+ n_2 +...+ n_i =n, то

$$D_в = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_в)^2 n_i .$$

15. Для характеристики розсіювання значень кількісної ознаки X генеральної сукупності навколо свого середнього генерального значення, вводять поняття генеральної дисперсії.

Генеральною дисперсією ($D_г$) називається середнє арифметичне квадратів відхилень значень ознаки X генеральної сукупності від його середнього генерального значення ($\overline{x}_г$):

$$D_г = \frac{1}{N}\sum_{i=1}^{n}(x_i - \overline{x}_u)^2 \text{ , якщо всі } x_i \text{ різні.}$$

Якщо варіанти x_i появляються відповідно N_1, N_2,..., N_i разів, де $N_1 + N_2 + \ldots + N_i = N$, то $D_г = \frac{1}{N}\sum_{i=1}^{n}(x_i - \overline{x}_г)^2 N_i$.

Вибіркова дисперсія $D_в$ є зсуненою оцінкою генеральної дисперсії $D_г$, оскільки математичне сподівання $M(D) = \frac{n-1}{n} D_в \neq D_г$.

16. *Генеральним (вибірковим) середнім квадратичним відхиленням (стандартом)* називається величина $\sigma_г = \sqrt{D_г}$ ($\sigma_в = \sqrt{D_в}$).

17. Незсуненою оцінкою генеральної дисперсії є *виправлена дисперсія*

$$S^2 = \frac{n}{n-1} D_в .$$

18. Незсуненою оцінкою генерального середнього квадратичного відхилення є *виправлене середнє квадратичне відхилення*

$$S = \sqrt{\frac{n}{n-1} D}$$

19. Але при $n \geq 30$ практично не розрізняють, де вибіркова, де виправлена дисперсія. Можна довести, що дисперсія від середнього вибіркового дорівнює дисперсії вибірки. Визначимо дисперсію вибіркової середньої:

$$D(\overline{x}) = D(\frac{1}{n}(x_1 + x_2 + \ldots + x_i))$$

$$= \frac{1}{n^2}(D(x_1) + D(x_2) + \ldots + D(x_i)) = \frac{1}{n^2} nD(x) = D(x)/n.$$

Висновок: $D(\overline{X}) = D(X)/n$.

20. Результати точкової оцінки можна записати так:

За статистичну оцінку параметра генеральної сукупності приймають її вибіркове значення $\overline{x}_в$ з середньою похибкою $\sigma_в$.

Дисперсія на практиці ототожнюється з похибкою, а середнє квадратичне відхилення вибірки – з середньою похибкою.

21. **Приклад 1**: З плодового дерева випадково зірвано 10 плодів, їх маси подані в таблиці. Опрацювати статистичні дані вибірки:

i	1	2	3	4	5	6	7	8	9	10
x	225	274	305	253	220	245	211	234	230	231
x-c	-25	24	55	3	-30	-5	-39	-16	-20	-19
$(x-c)^2$	625	576	3025	9	900	25	1521	256	400	361

C=250 ; $\overline{x} = 242.8$ (г) ; D = 707,96 (г^2) ; $\sigma_в$ =26,61 (г), S = 28,2 (г).

Отже, середня маса плода становить 242,8 г з похибкою 707,96 г^2, середньою похибкою 26,61 г.

Приклад 2. З генеральної сукупності зроблена вибірка:

x_i	2	5	7	10
n_i	16	12	8	14

Дати статистичну оцінку генеральній сукупності.

Розв'язування. Об'єм вибірки n=16+12+8+14=50,

$\overline{x}_в$ =(1/50) (2*16+5*12+7*8+10*14)=(1/50)*288=5,76;

$D_в$=(1/50)((2-5,76)^2 16+(5-5,6)^2 12+(7-5,6)^2 8+(10-5,6)^2 14)=144,73/50=2,895;

$$\sigma_в = \sqrt{2{,}895} \approx 1{,}70;$$

S^2=(50/49)*2,895=2,95; S=1,72.

22. *Властивості дисперсії.*

– Дисперсію можна обчислити за іншою формулою, яка більш зручна для обчислень: $D = \overline{x^2} - (\overline{x})^2$.

– Дисперсія постійної дорівнює нулю: D(C)=0.

– Постійний множник можна виносити за знак дисперсії, якщо попередньо піднести його до квадрату: $D(CX)=C^2 D(X)$.

– Дисперсія суми незалежних випадкових величин дорівнює сумі дисперсій доданків: $D(X_1 X_2 \ldots X_n) = D(X_1)+ D(X_2)+\ldots+ D(X_n)$.

– Якщо кожну ознаку змінити на певну величину, то дисперсія не зміниться: D(C+X)= D(X).

– Якщо кожну ознаку змінити у k разів (тобто ввести умовну варіанту $u=kx$), то дисперсія зміниться у k^2 разів. Тому, знайшовши дисперсію умовних варіант, треба розділити її на k^2: $D(X)=(1/k^2) D(U)$.

Питання для самоперевірки

1. Назвіть числові характеристики статистичного розподілу. Дайте означення цих характеристик.

2. Сформулюйте властивості числових характеристик статистичного розподілу.

3. Яка оцінка для математичного сподівання має властивість спроможності і незсуненості?

4. Яка оцінка для дисперсії має властивість має властивість спроможності і незсуненості?

ІНТЕРВАЛИ НАДІЙНОСТІ НОРМАЛЬНО РОЗПОДІЛЕНОЇ СТАТИСТИЧНОЇ ВЕЛИЧИНИ

1. Задачу інтервальної оцінки можна сформулювати так: за даними вибірки побудувати числовий інтервал, відносно якого із заздалегідь вибраною ймовірністю можна сказати, що всередині нього знаходиться параметр, який оцінюється. Інтервальна оцінка особливо необхідна при малій кількості спостережень, коли точкова оцінка малонадійна.

2. ***Інтервальною*** називається оцінка, яка визначається двома числами – кінцями інтервалу, який покриває параметр статистичної сукупності, який оцінюється.

3. Нехай θ оцінюваний параметр, θ^* – його оцінка, складена із X_1, X_2, ..., X_n. Нехай δ>0 – деяке число. Якщо виконується нерівність

$$\left|\theta-\theta^*\right|<\delta \text{, тобто } \quad -\delta<\theta-\theta^*<\delta, \quad \theta^*-\delta<\theta<\theta^*+\delta,$$

то кажуть, що інтервал $(\theta^*-\delta;\theta^*+\delta)$ покриває параметр θ. Число δ називається точністю оцінки θ^*.

Надійністю (надійним інтервалом) оцінки θ^* параметра θ для заданого числа $\delta>0$ називається ймовірність γ того, що інтервал $(\theta^*-\delta;\theta^*+\delta)$ покриє параметр θ. Але оскільки параметр θ нам невідомий, то ми називаємо ***надійністю*** γ уже обчисленої оцінки θ^* ймовірність того, що інтервал $(\theta^*-\delta;\theta^*+\delta)$, знайдений для довільної вибірки, покриє параметр θ. Отже, γ – міра нашої довіри обчисленої оцінки. Відзначимо, що після того, як за даними вибірки обчислена оцінка θ^*, то подія $\left\{\left|\theta^*-\theta\right|<\delta\right\}$ стає або достовірною або неможливою, тому що вказаний інтервал або покриває параметр θ, або ні. Отже, $(\theta^*-\delta;\theta^*+\delta)$ – надійний інтервал, γ – надійність, δ – точність оцінки, $\theta^*-\delta$ і $\theta^*+\delta$ – відповідно нижня та верхня межі (кінці) цього інтервалу.

4. Зрозуміло, що чим менше число δ, тим менша надійність γ. Надійність γ приймають рівною 0,95, або 0,99, або0,999. Це значить, що якщо зробити багато вибірок, то для 95% із них (якщо γ=0,95) обчислені інтервали надійності дійсно покриють θ.

5. Нехай $x_1, x_2, \dots, x_n$ – вибірка з генеральної сукупності X, для якої: a=M(X). Запишемо у вигляді таблиці теоретично обгрунтовані надійні інтервали для математичного сподівання **а** (за вибірковим середнім $\bar{x}$) і для середнього квадратичного відхилення σ (за вибірковим середнім квадратичним відхиленням S^2).

№ п/п	Параметри розподілу генеральної сукупності, який оцінюється	Тип розподілу генеральної сукупності	γ- надійні інтервали
1.	θ=**a**, σ – відоме	Нормальний	$\left(\bar{x}_в-\frac{t\sigma}{\sqrt{n}};\bar{x}_в+\frac{t\sigma}{\sqrt{n}}\right)$, де $\delta=\pm\frac{t\sigma}{\sqrt{n}}$, t знаходимо з рівності 2Φ(t)=γ, Φ(t) – табульована функція (табл.2 Додатків).
		Довільний (n≥30)	Φ(t)=γ/2.
2.	θ=**a**, σ – невідоме	Нормальний	$\left(\bar{x}_в-\frac{t_\gamma s}{\sqrt{n}};\bar{x}_в+\frac{t_\gamma s}{\sqrt{n}}\right)$, де t_γ= t(γ,n) – з табл.3.
		Довільний (n>30)	$\left(\bar{x}_в-\frac{ts}{\sqrt{n}};\bar{x}_в+\frac{ts}{\sqrt{n}}\right)$, де Φ(t)=γ/2 (табл. 2).
3.	θ=σ	Нормальний	(S-Sq; S+Sq), якщо q<1, (0;S+Sq), якщо q≥1, де q=q(γ;n) – з табл.4.

Отже, для оцінки істинного значення вимірюваної величини треба обчислити інтервал надійності виду 1 чи 2; для оцінки точності вимірювань (точності приладу) – інтервал виду 3.

Приклад 1. Ознака розподілена в генеральній сукупності нормально з відомим середнім квадратичним відхиленням σ=0,4. Знайти інтервал надійності для **а**, якщо γ =0,99, n =20, $\bar{x}$ =6,34.

Розв'язування. Для надійності γ =0,99 знаходимо за таблицею Додатків 2 для $\Phi(t)=\gamma/2=0,99/2=0,495$ значення $t=2,58$. Точність оцінки $\delta=\pm\frac{2,58\cdot 0,40}{\sqrt{20}}\approx 0,23$. Кінці (межі) інтервалу надійності 6,34-0,23=6,11 та 6,34+0,23=6,57. Отже, інтервал надійності (6,11; 6,57) покриває **а** з надійністю 0,99.

Приклад 2. Ознака X розподілена в генеральній сукупності нормально. Знайти надійний інтервал для $\bar{x}_г$ з надійністю γ = 0,99, n =20; $\bar{x}_в$=6,34; s =0,40.

Розв'язування. Для надійності γ =0,99 і n=20 знаходимо за таблицею Додатків 3 t_γ = 2,861. Отже, $\delta=\frac{2,861\cdot 0,40}{\sqrt{20}}\approx 0,26$. Тому кінці інтервалу надійності будуть 6,34-0,26=6,08; 6,34+0,26=6,60, отже, інтервал надійності (6,08; 6,60) покриває $\bar{x}_г$ з надійністю 0,99.

Приклад 3. Ознака X розподілена в генеральній сукупності нормально. Знайти надійний інтервал для $\sigma_г$ з надійністю γ=0,95, n =20; s=0,40.

Розв'язування. Для надійності γ =0,95 і n=20 знаходимо за таблицею Додатків 4 q=0,37, тоді sq=0,40·0,37≈0,15, тому кінці інтервальної надійності: 0,40-0,15=0,25; 0,40+0,15=0,55. Отже, інтервал надійності (0,25; 0,55) покриває $\sigma_г$ з надійністю 0,95.

Приклад 4. За даними вибіркового спостереження за якістю ґрунту у 9 господарствах провести точкову та інтервальні оцінки середньої якості ґрунту (в балах) в генеральній сукупності. Дані якості ґрунту подані в таблиці:

№ господарства	1	2	3	4	5	6	7	8	9
Якість ґрунту (x_i) в балах	70	79	80	71	77	74	67	77	75

Розв'язування. 1. Для визначення середньої якості ґрунту просумуємо x_i, а для обчислення дисперсії піднесемо їх до квадрату (x_i^2) і просумуємо. Дані запишемо у таблицю:

№ господарства	1	2	3	4	5	6	7	8	9	Всього
Якість ґрунту (x_i) в балах	70	79	80	71	77	74	67	77	75	670
Квадрат якості (x_i^2)	4900	6241	6400	5041	5929	5476	4489	5939	5625	50030 50040

6. Визначимо середнє значення якості ґрунту за вибірковою середньою: $\bar{x}_в=\frac{1}{n}\sum x_i=\frac{670}{9}=74,4$ (бала).

7. Обчислимо вибіркову дисперсію:

$$Д_в=\frac{\sum x_i^2}{n}-(\bar{x})^2=\frac{50030}{9}-74,4^2=23,53\,.$$

8. Обчислимо виправлену дисперсію:

$$s^2=\frac{n}{n-1}D_в=23,53\cdot\frac{9}{9-1}=26,47\text{, або}$$

$$s^2\frac{\sum x_i-n\cdot(\bar{x})^2}{n-1}=\frac{50030-9\cdot 74,4^2}{9-1}=26,47\,.$$

9. Знайдемо середню похибку вибіркової середньої:

$$\mu=\sqrt{\frac{s^2}{n}}=\sqrt{\frac{26,47}{9}}=\sqrt{2,94}\approx 1,715\approx 1,7\text{ бала.}$$

10. Проведемо точкову оцінку середньої якості ґрунту в генеральній сукупності: $\bar{x}=\bar{x}_0$=74,4 бала при середній похибці $\mu_{\bar{x}}$=1,7 бала, тобто точкова оцінка генеральної середньої може бути записана як $\bar{x}\pm\mu_{\bar{x}}$=74,4±1,7 бала. Це значить, що $\bar{x}$=74,4 бала є оцінкою генеральної середньої з похибкою, що дорівнює 1,7 бала.

11. Для проведення інтервальної оцінки генеральної сукупності і побудови інтервала надійності визначимо межову похибку вибіркової середньої для малих вибірок. Для цього за таблицею 4 Додатків при заданому рівні ймовірності γ=0,95 та n=9 маємо: t_γ=2,31.

12. Межова похибка вибіркової середньої δ=2,31·1,7=3,927≈3,93 бала.

13. Проведемо інтервальну оцінку середньої якості ґрунту в генеральній сукупності: $\bar{x}_0=\bar{x}\pm\delta$=74,4±3,93 бала або 70,47< $\bar{x}_0$<78,33.

Отже, з надійністю γ=0,95 можна стверджувати, що середня якість грунту знаходиться в інтервалі (70,47; 78,33) балів.

Питання для самоперевірки

1. Що називається надійним інтервалом і надійною ймовірністю (надійністю)?
2. Як будується інтервал надійності для математичного сподівання випадкової величини, яка розподілена за нормальним законом?
3. Чим відрізняються точкова та інтервальна оцінки?

ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

1. Перевірка статистичних гіпотез тісно пов'язана з теорією оцінки параметрів розподілу. В практичній і науковій діяльності часто для доведення справедливості того чи іншого факту висувають гіпотезу, яка може бути перевірена на основі даних вибіркового спостереження.

2. Під ***гіпотезою*** будемо розуміти в широкому смислі слова деяке наукове припущення про властивості явищ, що вивчаються, яке вимагає перевірки і доведення

3. Під ***статистичною гіпотезою*** (Н) розуміють припущення відносно параметрів або форми розподілу генеральної сукупності, яке перевіряються на основі даних вибіркового спостереження.

З цього означення випливає, що статистична гіпотеза може відноситись або до законів розподілу, або до окремих параметрів розподілу.

4. Прикладом статистичних гіпотез можуть бути припущення про те, що жива маса телят в господарстві якогось села або району підпорядковується закону нормального розподілу; або середня урожайність картоплі одного сорту перевищує середній урожай картоплі іншого сорту; середні розміри деталей, що виготовляються на однотипних станках, однакові тощо.

5. У ході перевірки статистичної гіпотези необхідно перевірити чи узгоджуються дані спостереження з висунутою гіпотезою, чи можна відмінності між гіпотезою і даними спостереження віднести до випадкових (неістотних, несуттєвих) або ж ці відмінності викликані впливом якихось систематично діючих причин. В результаті перевірки гіпотеза або *спростовується (відхиляється)*, або *приймається*.

6. Найбільшими часто перевіряється припущення про те, що отримана за вибіркою величина досить мало відрізняється від гіпотетичної (теоретично припущеної) або встановленої величини генеральної сукупності.

Для перевірки цього положення висувається гіпотеза про те, що справжня різниця між фактичними і гіпотетичними показниками дорівнює нулю. Тому гіпотезу, що перевіряється, називають *нульовою (основною, робочою)* (H_0).

7. В кожному випадку нульовій гіпотезі Н0 протиставляють альтернативну (конкуруючу) (H_a чи H_1), яка заперечує H_0.

Наприклад, якщо Но полягає у припущенні, що середня врожайність зернових культур в генеральній сукупності дорівнює 35 ц/га, то H_a може бути така: середня врожайність зернових культур не дорівнює 35 ц/га. Записують:

H_0: $\bar{x}=35$; H_a: $\bar{x}\neq 35$.

8. По відношенню до H_0 може бути вказана довільна кількість H_a. Зміст H_0 та H_a залежить від характеру вибіркової величини; конкретних

задач, які розв'язуються за допомогою перевірки статистичних гіпотез. Формально будь-яка з двох гіпотез, що протиставляються може бути нульовою. Але на практиці наслідки можуть бути різними у зв'язку з різним характером помилок отриманого в результаті неправильного прийняття або відхилення гіпотези H_0. Тому вибір H_0 вимагає спеціального обґрунтування.

9. По формі побудови гіпотези бувають прості і складні.

Простою називається гіпотеза, яка містить тільки одне припущення (H_0: $\sigma=3$). У противному разі гіпотеза називається *складною* (H_0: $\sigma>3$ *або* H_0: $\sigma<3$, оскільки σ може бути довільним числом, що задовольняє умові, тобто ці гіпотези складаються з декількох простих гіпотез).

10. перевірки H_0 і прийняття висновку про сумісність вибіркових даних з висунутою гіпотезою використовуються спеціальні *статистичні критерії*, які являють собою список правил, за якими гіпотеза, що перевіряється, приймається або відхиляється.

11. кожного виду гіпотез, що перевіряються, розроблені спеціальні критерії: t-критерій нормального розподілу і розподілу Стьюдента, F-критерій Фішера-Снедекора, χ^2-критерій Пірсона.

12. У множині можливих значень вибраного критерію (деякої випадкової величини k, яка служить для перевірки H_0), можна виділити дві підмножини, які не перетинаються: *критична область* (ті значення критерію, при яких H_0 гіпотеза відхиляється) та *область допустимих значень чи область прийняття гіпотези* (сукупність значень критерію, що використовується, при яких H_0 приймається). Точки, які відокремлюють критичну область від області допустимих значень, називаються *критичними точками (межами)*.

13. Критичні області бувають односторонні і двосторонні. *Односторонньою* називають правосторонню (лівосторонню) критичну область, яка виражається нерівністю $k>a_{крит}$. ($k<a_{крит}$), де $a_{крит}$ – додатне (від'ємне) число. *Двустороння* критична область визначається нерівністю $|k|>a_{крит}$, *де* $a_{крит}>0$.

ОСНОВНИЙ ПРИНЦИП ПЕРЕВІРКИ СТАТИСТИЧНИХ ГІПОТЕЗ

1. Якщо фактичне спостережуване значення критерію належить критичній області, то H_0 відхиляють; якщо фактичне значення належить області прийняття, то H_0 приймають.

2. Для кожного критерію складені відповідні спеціальні таблиці, за якими знаходять його табличне значення (критичні точки). Знайдені табличні значення співставляють з його фактичним значенням.

Якщо фактичне значення критерію, розраховане за даними вибірками, більше від табличного значення, то H_0 має бути *відхиленою*, і прийнятою H_a.

Якщо фактичне значення критерію виявляється меншим або рівним табличним значенням, то робиться висновок про узгодження H_0 та спостережуваних даних, тобто основ для відмови від H_0 немає і тому вона може бути прийнята.

3. Статистичні критерії, які використовуються для перевірки статистичної гіпотези, поділяються на 2 види: параметричні; непараметричні.

Параметричними називаються критерії, які базуються на тому припущенні, що розподіл випадкової величини в сукупності підпорядковується деякому відомому закону (нормальному, Пуансона та ін.). До таких критеріїв відносяться критерії t, F, χ^2.

Непараметричними (порядковими) називаються ті критерії, використання яких не пов'язане зі знанням розподілу випадкової величини (критерій знаків, Вілкоксона, Уайта, Манна-Уітни та ін.).

Параметричний критерій більш ефективний в порівнянні з непараметричним критерієм.

4. *Перевірка гіпотези про нормальний розподіл генеральної сукупності за критерієм χ^2 (хі-квадрат) Пірсона.*

Нехай задано емпіричний розподіл (у вигляді послідовності рівновіддалених варіант і відповідних їм частот).

x_i	x_1	x_2	...	x_N
n_i	n_1	n_2	...	n_N

Потрібно за критерієм Пірсона перевірити гіпотезу про те, що генеральна сукупність X розподілена нормально (тобто визначається двома параметрами: $M=a$ та середнім квадратичним відхиленням σ).

Правило. Для того, щоб при заданому рівні значимості критерію α перевірити гіпотезу про нормальний розподіл генеральної сукупності треба:

1. Обчислити безпосередньо (при малій кількості спостережень) або спрощеним методом (при великій кількості спостережень) $\overline{x}_в$ та $\sigma_в$.

2. Обчислити теоретичні частоти $n_i' = \frac{nh}{\sigma}\cdot\varphi(u_i)$, де n – об'єм вибірки, h – крок (різниця між двома сусідніми варіантами), $u_i = \frac{x_i - \overline{x}_в}{\sigma_в}$, а значення функції $\varphi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$ – знаходять за табл. 1 Додатків.

3. Порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона. Для цього: а) скласти розрахункову таблицю, за якою знайти спостережуване (фактичне, емпіричне) значення критерію

$$\chi^2_c = \sum \frac{(n_i - n_i')^2}{n_i'};$$

б) за таблицею критичних точок розподілу χ^2 за даним рівнем значимості α і числом ступенів волі $k=s-3$ (s – це число груп вибірки) знаходять критичну (табличну) точку $\chi^2_{кр}(\alpha;\ k)$ правосторонньої критичної області. Число α – наперед задана достатньо мала величина – та мінімальна ймовірність, починаючи з якої можна признати подію практично неможливою. Найчастіше беруть $\alpha=0,05$; 0,01; 0,001 та ін. Це означає, що у середньому в 5-ти (1, 1/10, ...) випадках із 100 є риск відхилити правильну гіпотезу H_0.

4. Якщо $\chi^2_c \le \chi^2_{кр}$, то нульова гіпотеза H_0 приймається. Тобто емпіричні і теоретичні частоти відрізняються дуже мало (випадково).

Якщо $\chi^2_c > \chi^2_{кр}$, то H_0 відхиляють. Тобто емпіричні і теоретичні частоти відрізняються значно.

Фактичне значення χ^2 можна розраховувати за іншою формулою:

$$\chi^2 = \sum_{i=1}^{n} \frac{n_i^2}{n_i'} - n.$$

5. Критерій Пірсона використовується для розв'язання ряду задач. Наприклад: при перевірці задач про узгоджуваність (згоду) вибіркового і теоретичного розподілу; про незалежність розподілів, про однорідність розподілів тощо. Відповідно до цих задач критерій χ^2 називається критерієм згоди, незалежності, однорідності.

6. Застосування χ^2 – критерію вимагає дотримання деяких умов. Найважливішими з них є: 1) Об'єм вибірки має бути достатньо великим (при n<50 потужність критерію знижується); 2) Чисельність окремих інтервалів (класів) має містити не менше 5 одиниць; в противному випадку малочисельні інтервали об'єднують, при цьому об'єднують і відповідні частоти. Якщо відбулось чергове об'єднання частот, то при знаходженні числа ступенів волі треба за s брати кількість груп вибірки, що залишились після об'єднання частот.

Приклад 1. Використовуючи критерій Пірсона для рівня значимості $\alpha=0,05$ перевірити, чи узгоджується гіпотеза про нормальний розподіл генеральної сукупності X з емпіричним розпорядком вибірки об'ємом n=200:

x_i	5	7	9	11	13	15	17	19	21
n_i	15	26	25	30	26	21	24	20	13

Розв'язування:

1. Знайдемо середнє вибіркове $\overline{x}_в$=12,63 і середнє квадратичне відхилення σ= 4,695.

2. Обчислимо теоретичні частоти, враховуючи, що n=200, σ=4,695, h=2, за формулою

$$n_i' = \frac{nh}{\sigma} \cdot \varphi(u_i) = \frac{200 \cdot 2}{4,695} \cdot \varphi(u) = 85,2 \cdot \varphi(u).$$

Складемо розрахункову таблицю , для якої значення φ(u) знайдемо за таблицею 1 Додатків:

i	$x_ш$	u_i	$\varphi(u_i)$	$n_i'=85,2\varphi(u_i)$
1	5	-1,62	0,1074	9,1
2	7	-1,20	0,1942	16,5
3	9	-0,77	0,2966	25,3
4	11	-0,35	0,3752	32,0
5	13	0,08	0,3977	33,9
6	15	0,51	0,3503	29,8
7	17	0,93	0,2589	22,0
8	19	1,36	0,1582	13,5
9	21	1,78	0,0818	7,0

Порівняємо емпіричні та теоретичні частоти.

А) Складемо розрахункову таблицю 2, з якої знайдемо спостережуване значення критерію χ^2:

i	n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$(n_i - n_i')^2/n_i'$
1	15	9,1	5,9	34,81	3,8
2	26	16,5	9,5	90,25	5,5
3	25	25,3	-0,3	0,09	0,0
4	30	32,0	-2,0	4,00	0,1
5	26	33,9	-7,9	62,41	1,8
6	21	29,8	-8,8	77,44	2,6
7	24	22,0	2,0	4,0	0,2
8	30	13,5	6,5	42,25	3,1
9	13	7,0	6,0	36,00	5,1
Всього	200	-	-	-	χ^2_c=22,2

б) За таблицею 5 критичних точок розподілу χ^2 за рівнем значимості α=0,05 і числом ступенів волі k=9-3=6 знаходимо критичну точку правосторонньої критичної області $\chi^2_{крит}$.(0,05; 6)=12,6.

4. Оскільки $\chi^2_c > \chi^2_{кр}$, тобто 22,2>12,6, то гіпотеза про нормальний розподіл генеральної сукупності відхиляється, тобто емпіричні і теоретичні частоти відрізняються значно.

Питання для самоперевірки

1. Сформулюйте задачу статистичної перевірки гіпотез.
2. Наведіть приклади задач на перевірку гіпотез.

3. В чому полягає ідея застосування критеріїв згоди при розв'язуванні задач про узгодження теоретичного та емпіричного розподілів?

ЕЛЕМЕНТИ ТЕОРІЇ КОРЕЛЯЦІЇ

1. Часто доводиться мати справу з більш складною залежністю, ніж функціональна. Якщо кожному значенню однієї величини відповідає множина можливих значень іншої, то такі залежності належать до *кореляційних залежностей.*

2. ***Статистичною*** називають залежність, при якій зміна однієї із величин призводить до зміни розподілу іншої.

3. Якщо при зміні однієї із величин змінюється середнє значення другої, то статистичну залежність називають ***кореляційною***.

4. Нехай вивчається зв'язок між випадковими величинами Y та X. Нехай кожному значенню X відповідає декілька значень Y: $y_1, y_2, \ldots, y_k$. Середнє арифметичне цих чисел є число $\overline{y}_x = \frac{y_1 \cdot y_2 \cdot \ldots \cdot y_k}{k}$, яке називають *умовним середнім*, що відповідає значенню X=x.

5. *Кореляційною залежністю* Y від X називають функціональну залежність умовної середньої $\overline{y}_x$ від x: $\overline{y}_x$=*f(x)*. Це рівняння називають ***рівнянням регресії*** Y на X; функцію *f(x)* називають *регресією* Y на X, а її графік – лінією регресії Y на X. Аналогічно означаються умовна середня $\overline{x}_y$ та кореляційна залежність X від Y: $\overline{x}_y = ky + m$.

6. Математична статистика розв'язує дві основні задачі теорії кореляції: а) встановити форму кореляційного зв'язку, тобто вид функції регресії (лінійна, квадратична, показникова і т.д.). Якщо при цьому функції (лінії) регресії Y на X і X на Y – прямі, то кореляцію називають *лінійною*. б) оцінити тісноту (силу) зв'язку, яка оцінюється за величиною розсіювання значень навколо умовної середньої. Чим більша розсіюваність, тим слабша залежність.

7. Рівняння регресії Y на X при лінійному кореляційному зв'язку можна записати у вигляді $\overline{y}_x = ax + b$. Кутовий коефіцієнт *а* називають вибірковим коефіцієнтом регресії і позначають через ρ_{yx}. Коефіцієнти *а* та *b* за допомогою методу найменших квадратів знаходять із системи:

$$\begin{cases} b\sum x_i + a\sum x_i^2 = \sum x_i y_i, \\ nb + a\sum x_i = \sum y_i. \end{cases}$$

Приклад 1. Знайти вибіркове рівняння прямої регресії Y на X за даними $n=5$ спостережень:

x_i	1,00	1,50	3,00	4,50	5,00
y_i	1,25	1,40	1,50	1,75	2,25

Розв'язування. Для підрахунку коефіцієнтів системи складемо розрахункову таблицю:

x_i	y_i	x_i^2	x_iy_i
1,00	1,25	1,00	1,250
1,50	1,40	2,25	2,100
3,00	1,50	9,00	4,500
4,50	1,75	20,25	4,875
5,00	2,25	25,00	11,250
$\sum x_i$ =15	$\sum y_i$ =8,15	$\sum x^2_i$ =57,50	$\sum x_iy_i$ =26,975

Із системи отримаємо значення a та b:

$$a=\frac{n\sum xy-\sum x\cdot\sum y}{n\sum x^2-(\sum x^2)},\quad \mathrm{b}=\frac{\sum \mathrm{x}^2\cdot\sum y-\sum x\cdot\sum xy}{n\sum x^2-(\sum x^2)}.$$

Підставивши значення сум із таблиці, отримаємо:

$$a=\frac{5\cdot 26{,}975-15\cdot 8{,}15}{5\cdot 57{,}5-15^2}\approx 0{,}202,\quad \mathrm{b}=\frac{57{,}5\cdot 8{,}15-15\cdot 26{,}975}{62{,}5}\approx 1{,}024.$$

Тому шукане рівняння регресії Y на X має вид: $\overline{y_x}=0{,}202x+1{,}024$.

9. Вибіркове рівняння прямої лінії регресії (Y на X) прийнято записувати ще в іншій формі: $\overline{y_x}-\overline{y}=\rho_в\dfrac{\sigma_y}{\sigma_x}(x-\overline{x})$, де $\overline{y_x}$ – умовна середня, $\overline{x}$ та $\overline{y}$ – вибіркова середня ознак X та Y, σx та σy – вибіркове середнє квадратичне відхилення, $\rho_в$ – вибірковий коефіцієнт кореляції, який обчислюється за формулою: $\rho_в=\dfrac{\sum n_{xy}\cdot x\cdot y-n\cdot\overline{x}\cdot\overline{y}}{n\sigma_x\sigma_y}$.

Приклад 2. Знайти вибіркове рівняння прямої лінії регресії Y на X за даними кореляційної таблиці:

Y	X					n_y
	20	25	30	35	40	
16	4	6				10
26		8	10			18
36			32	3	9	44
46			4	12	6	22
				1	5	6
n_x	4	14	46	16	20	n=100

Розв'язування. 1. Складемо кореляційну таблицю в умовних варіантах, вибравши за "хибні" нулі C_1=30 та C_2=36.

v	u					n_v
	-2	-1	0	1	2	
-2	4	6				10
-1		8	10			18
0			32	3	9	44
1			4	12	6	22
2				1	5	6
n_u	4	14	46	16	20	n=100

3. Знайдемо $\overline{u}$ та $\overline{v}$: $\overline{u}=(\Sigma n_u\cdot u)/\ n=$
$=(4\cdot(-2)+14\cdot(-1)+46\cdot 0+16\cdot 1+20\cdot 2)/100=0{,}34$;
$\overline{v}=(\Sigma n_v\cdot v)/\ n)=(10\cdot(-2)+18\cdot(-1)+44\cdot 0+22\cdot 1+6\cdot 2)/100=-0{,}04$.

4. Знайдемо допоміжні величини $\overline{u}^2=(0.34)^2$ *та* $\overline{v}^2=(-0.04)^2$;

$$\sigma_u=\sqrt{\overline{u^2}-(\overline{u})^2}=\sqrt{1.26-0.34^2}\approx 10{,}7;\ \sigma_v=\sqrt{1.04-0.04^2}\approx 1.02$$

5. Обчислимо:

$$\sum n_{uv}\cdot uv=4\cdot(-2)\cdot(-2)+6\cdot(-2)\cdot(-1)+8\cdot(-1)\cdot(-1)+10\cdot(-1)\cdot 0+$$
$$+32\cdot 0\cdot 0+4\cdot 1\cdot 0+3\cdot 0\cdot 1+12\cdot 1\cdot 1+1\cdot 2\cdot 1+9\cdot 0\cdot 2+6\cdot 1\cdot 2+5\cdot 2\cdot 2=82.$$

6. Знайдемо шуканий вибірковий коефіцієнт кореляції:

$$\rho_в=\frac{\sum n_{uv}\cdot uv-n\overline{uv}}{n\sigma_u\sigma_v}=\frac{82-100\cdot 0.34\cdot(-0.04)}{100\cdot 1.07\cdot 1.02}=0.76.$$

7. Знайдемо кроки h_1 та h_2 (різниці між двома сусідніми варіантами): $h_1=25-20=5$; $h_2=26-16=10$.

8. Знайдемо $\overline{x}$ та $\overline{y}$, враховуючи, що $C_1=30$, $C_2=36$:
$\overline{x}=\overline{u}\cdot h_1+C_1=0{,}34\cdot 5+30=31{,}70$; $\overline{y}=\overline{v}\cdot h_2+C_2=(-0{,}04)\cdot 10+36=35{,}60$.

9. Обчислимо: $\sigma_x=h_1\cdot\sigma_u=5\cdot 1.07=5.35$; $\sigma_y=h_1\cdot\sigma_v=10\cdot 1{,}02=10{,}2$.

Тоді шукане рівняння прямої лінії регресії Y на X буде:

$$\overline{y_x}-35{,}60=0{,}76\cdot\frac{10{,}2}{5{,}35}\cdot(x-31{,}70),\ \text{або}\ \overline{y_x}=1{,}45x-10{,}36.$$

КОНТРОЛЬНА РОБОТА

1 Завдання. Знайти: а) вибіркову середню; б) вибіркову дисперсію; в) вибіркове середнє квадратичне відхилення за даним статистичним розподілом вибірки (в першому рядку вказані вибіркові варіанти x_i, у другому – відповідні частоти n_i кількісної ознаки X).

1. x_i	105	110	115	120	125	130	135
n_i	4	6	10	40	20	12	8
2. x_i	12,5	13,0	13,5	14,0	14,5	15,0	15,5
n_i	5	15	40	25	8	4	3
3. x_i	10,2	10,9	11,6	12,3	13,0	13,7	14,4
n_i	8	10	60	12	5	3	2
4. x_i	45	50	55	60	65	70	75
n_i	4	6	10	40	20	12	8
5 x_i	110	115	120	125	130	135	140
n_i	5	10	30	25	15	10	5
6. x_i	12,4	16,4	20,4	24,4	28,4	32,4	36,4
n_i	5	15	40	25	8	4	3
7. x_i	26	32	38	44	50	56	62
n_i	5	15	40	25	8	4	3
8. x_i	10,6	15,6	20,6	25,6	30,6	35,6	40,6
n_i	8	10	60	12	5	3	2
9. x_i	100	110	120	130	140	150	160
n_i	4	6	10	40	20	12	8
10. x_i	130	140	150	160	170	180	190
n_i	5	10	30	25	15	10	5

2 Завдання. Знайти інтервали надійності для оцінки математичного сподівання **a** нормального розподілу з надійністю 0,95, знаючи вибіркову середню $\bar{x}$,, об'єм вибірки *n* і середнє квадратичне відхилення σ.

1.	$\bar{x}=75,17;$	$\sigma=6;$	$n=36.$
2.	$\bar{x}=75,16;$	$\sigma=7;$	$n=49,$
3.	$\bar{x}=75,15;$	$\sigma=8;$	$n=64.$
4.	$\bar{x}=75,14;$	$\sigma=9;$	$n=81.$
5.	$\bar{x}=75,13;$	$\sigma=10;$	$n=100.$
6	$\bar{x}=75,12;$	$\sigma=11;$	$n=121.$
7.	$\bar{x}=75,11;$	$\sigma=12;$	$n=144.$
8.	$\bar{x}=75,10;$	$\sigma=13;$	$n=169.$
9.	$\bar{x}=75,09;$	$\sigma=14;$	$n=196$
10.	$\bar{x}=75,08;$	$\sigma=15;$	$n=225.$

3 Завдання. Задано вибіркові середні $\bar{x}$ та $\bar{y}$, знайдені за вибірками об'ємів n = 60 і m = 50, вибрані із нормальних генеральних сукупностей X та Y з відомими дисперсіями D(X) і D(Y). Треба при рівні значимості α= 0,05 перевірити нульову гіпотезу Ho: M(X) = M(Y) при конкуруючій гіпотезі Ha: M(X) ≠ M(Y), тобто треба встановити, значно чи незначно відрізняються вибіркові середні.

1.	$\bar{x}=638$	$\bar{y}=625$	$D(X)=30$	$D(Y)=25.$
2.	$\bar{x}=130$	$\bar{y}=125$	$D(X)=60$	$D(Y)=150.$
3.	$\bar{x}=260$	$\bar{y}=250$	$D(X)=90$	$D(Y)=125.$
4.	$\bar{x}=390$	$\bar{y}=375$	$D(X)=180$	$D(Y)=50.$
5.	$\bar{x}=520$	$\bar{y}=500$	$D(X)=72$	$D(Y)=140.$
6.	$\bar{x}=650$	$\bar{y}=640$	$D(X)=78$	$D(Y)=135.$
7.	$\bar{x}=780$	$\bar{y}=785$	$D(X)=120$	$D(Y)=100.$
8.	$\bar{x}=910$	$\bar{y}=885$	$D(X)=180$	$D(Y)=50.$
9.	$\bar{x}=1000$	$\bar{y}=990$	$D(X)=84$	$D(Y)=130.$
10.	$\bar{x}=13.8$	$\bar{y}=17.1$	$D(X)=96$	$D(Y)=120.$

4 Завдання. Треба при рівні значимості α=0,05 перевірити за критерієм згоди Пірсона гіпотезу про нормальний розподіл генеральної сукупності, якщо відомі емпіричні частоти n_i та теоретичні частоти n_i'.

1. n_i	6	10	20	27	19	12	6
n_i'	5	14	16	25	21	12	7.
2. n_i	5	12	19	27	20	10	7
n_i'	6	13	14	27	22	12	6.
3. n_i	8	10	18	27	17	11	9
n_i'	4	15	16	25	21	12	7
4. n_i	5	11	22	25	21	11	5
n_i'	5	13	17	25	21	12	7.
5. n_i	7	12	16	25	21	11	7
n_i'	5	12	18	29	20	10	6.
6. n_i	8	12	16	27	19	12	6
n_i'	5	17	13	25	21	12	7.
7. n_i	6	15	16	26	19	12	6
ni'	5	17	13	25	21	12	7
8. n_i	6	13	21	23	19	12	6
ni'	5	14	16	25	21	13	6.

9. n_i 5 11 20 27 19 12 6
n_i' 6 13 16 25 20 13 7.

10. n_i 6 9 21 27 20 11 6
n_i' 5 13 17 25 21 12 7.

5 Завдання. Знайти вибіркове рівняння прямої $\overline{y}_x - \overline{y} = \rho_в \frac{\sigma_y}{\sigma_x}(x - \overline{x})$ регресії Y на X за даною кореляційною таблицею.

1.

Y	X						n_y
	10	15	20	25	30	35	
45	2	4	-	-	-	-	6
55	-	3	5	-	-	-	8
65	-	-	5	35	5	-	45
75	-	-	2	8	17	-	27
85	-	-	-	4	7	3	14
n_x	2	7	12	47	29	3	n=100

2.

Y	X						n_y
	10	15	20	25	30	35	
40	2	4	-	-	-	-	6
50	-	3	7	-	-	-	10
60	-	-	5	30	10	-	45
70	-	-	7	10	8	-	25
80	-	-	-	5	6	3	14
n_x	2	7	19	45	24	3	n=100

3.

Y	X						n_y
	15	20	25	30	35	40	
15	4	1	-	-	-	-	5
25	-	6	4	-	-	-	10
35	-	-	2	50	2	-	54
45	-	-	1	9	7	-	17
55	-	-	-	4	3	7	14
n_x	4	7	7	63	12	7	n=100

4.

Y	X						n_y
	2	7	12	17	22	27	
100	1	5	-	-	-	-	6
120	-	5	3	-	-	-	8
130	-	-	3	40	12	-	55
140	-	-	2	10	5	-	17
150	-	-	-	3	4	7	14
n_x	1	10	8	"53	21	7	n=100

5.

Y	X						n_y
	5	10	15	20	25	30	
10	3	5	-	-	-	-	8
20	-	4	4	-	-	-	8
30	-	-	7	35	8	-	50
40	-	-	2	10	8	-	20
50	-	-	-	5	6	3	14
n_x	3	10	21	47	15	4	n=100

6.

Y	X						n_y
	12	17	22	27	32	37	
25	2	4	-	-	-	-	6
35	-	6	3	-	-	-	9
45	-	-	6	35	4	-	45
55	-	-	2	8	6	-	16
65	-	-	-	14	7	3	24
n_x	2	10	11	57	17	3	n=100

7.

Y	X						n_y
	15	20	25	30	35	40	
25	3	4	-	-	-	-	7
35	-	6	3	-	-	-	9
45	-	-	6	35	2	-	43
55	-	-	12	8	6	-	26
65	-	-	-	4	7	4	15
n_x	3	10	21	47	15	4	n=100

8.

Y	X						n_y
	4	9	14	19	24	29	
30	3	3	-	-	-	-	6
40	-	5	4	-	-	-	9
50	-	-	40	2	8	-	50
60	-	-	5	10	6	-	21
70	-	-	-	4	7	3	14
n_x	3	9	49	16	21	3	n=100

9.

Y	X						n_y
	5	10	15	20	25	30	
30	2	6	-	-	-	-	8
40	-	5	3	-	-	-	8
50	-	-	7	40	2	-	49
60	-	-	4	9	6	-	19
70	-	-	-	4	7	5	16
n_x	2	11	14	53	15	5	n=100

10.

Y	X						n_y
	10	15	20	25	30	35	
20	5	1	-	-	-	-	6
30	•	6	2	-	-	-	8
40	-	-	5	40	5	-	50
50	-	-	2	8	7	-	17
60	-	-	-	4	7	8	19
n_x	5	7	9	52	19	8	n=100

ДОДАТКИ

Додаток 1

Таблица значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Додаток 2

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

x	Φ (x)	x	Φ (x)	x	Φ (x)	x	Φ (x)
0,00	0,0000	0,49	0,1879	0,98	0,3365	1,47	0,4292
0,01	0,0040	0,50	0,1915	0,99	0,3389	1,48	0,4306
0,02	0,0080	0,51	0,1950	1,00	0,3413	1,49	0,4319
0,03	0,0120	0,52	0,1985	1,01	0,3438	1,50	0,4332
0,04	0,0160	0,53	0,2019	1,02	0,3461	1,51	0,4345
0,05	0,0199	0,54	0,2054	1,03	0,3485	1,52	0,4357
0,06	0,0239	0,55	0,2088	1,04	0,3508	1,53	0,4370
0,07	0,0279	0,56	0,2123	1,05	0,3531	1,54	0,4382
0,08	0,0319	0,57	0,2157	1,06	0,3554	1,55	0,4394
0,09	0,0359	0,58	0,2190	1,07	0,3577	1,56	0,4406
0,10	0,0398	0,59	0,2224	1,08	0,3599	1,57	0,4418
0,11	0,0438	0,60	0,2257	1,09	0,3621	1,58	0,4429
0,12	0,0478	0,61	0,2291	1,10	0,3643	1,59	0,4441
0,13	0,0517	0,62	0,2324	1,11	0,3665	1,60	0,4452
0,14	0,0557	0,63	0,2357	1,12	0,3686	1,61	0,4463
0,15	0,0596	0,64	0,2389	1,13	0,3708	1,62	0,4474
0,16	0,0636	0,65	0,2422	1,14	0,3729	1,63	0,4484
0,17	0,0675	0,66	0,2454	1,15	0,3749	1,64	0,4495
0,18	0,0714	0,67	0,2486	1,16	0,3770	1,65	0,4505
0,19	0,0753	0,68	0,2517	1,17	0,3790	1,66	0,4515
0,20	0,0793	0,69	0,2549	1,18	0,3810	1,67	0,4525
0,21	0,0832	0,70	0,2580	1,19	0,3830	1,68	0,4535
0,22	0,0871	0,71	0,2611	1,20	0,3949	1,69	0,4545
0,23	0,0910	0,72	0,2642	1,21	0,3869	1,70	0,4554
0,24	0,0948	0,73	0,2673	1,22	0,3888	1,71	0,4564
0,25	0,0987	0,74	0,2703	1,23	0,3907	1,72	0,4573
0,26	0,1026	0,75	0,2734	1,24	0,3925	1,73	0,4582
0,27	0,1064	0,76	0,2764	1,25	0,3944	1,74	0,4591
0,28	0,1103	0,77	0,2794	1,26	0,3962	1,75	0,4599
0,29	0,1141	0,78	0,2823	1,27	0,3980	1,76	0,4608
0,30	0,1179	0,79	0,2852	1,28	0,3997	1,77	0,4616
0,31	0,1217	0,80	0,2881	1,29	0,4015	1,78	0,4625
0,32	0,1255	0,81	0,2910	1,30	0,4032	1,79	0,4633
0,33	0,1293	0,82	0,2939	1,31	0,4049	1,80	0,4641
0,34	0,1331	0,83	0,2967	1,32	0,4066	1,81	0,4649
0,35	0,1368	0,84	0,2995	1,33	0,4082	1,82	0,4656
0,36	0,1406	0,85	0,3023	1,34	0,4099	1,83	0,4664
0,37	0,1443	0,86	0,3051	1,35	0,4115	1,84	0,4671
0,38	0,1480	0,87	0,3078	1,36	0,4131	1,85	0,4678
0,39	0,1517	0,88	0,3106	1,37	0,4147	1,86	0,4686
0,40	0,1554	0,89	0,3133	1,38	0,4162	1,87	0,4693
0,41	0,1591	0,90	0,3159	1,39	0,4177	1,88	0,4699
0,42	0,1628	0,91	0,3186	1,40	0,4192	1,89	0,4706
0,43	0,1664	0,92	0,3212	1,41	0,4207	1,90	0,4713
0,44	0,1700	0,93	0,3238	1,42	0,4222	1,91	0,4719
0,45	0,1736	0,94	0,3264	1,43	0,4236	1,92	0,4726
0,46	0,1772	0,95	0,3289	1,44	0,4251	1,93	0,4732
0,47	0,1808	0,96	0,3315	1,45	0,4265	1,94	0,4738
0,48	0,1844	0,97	0,3340	1,46	0,4279	1,95	0,4744

Продовження додатку 2

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,96	0,4750	2,22	0,4868	2,54	0,4945	2,84	0,4977
1,97	0,4756	2,24	0,4875	2,56	0,4948	2,86	0,4979
1,98	0,4761	2,26	0,4881	2,58	0,4951	2,88	0,4980
1,99	0,4767	2,28	0,4887	2,60	0,4953	2,90	0,4981
2,00	0,4772	2,30	0,4893	2,62	0,4956	2,92	0,4982
2,02	0,4783	2,32	0,4898	2,64	0,4959	2,94	0,4984
2,04	0,4793	2,34	0,4904	2,66	0,4961	2,96	0,4985
2,06	0,4803	2,36	0,4909	2,68	0,4963	2,98	0,4986
2,08	0,4812	2,38	0,4913	2,70	0,4965	3,00	0,49865
2,10	0,4821	2,40	0,4918	2,72	0,4967	3,20	0,49931
2,12	0,4830	2,42	0,4922	2,74	0,4969	3,40	0,49966
2,14	0,4838	2,44	0,4927	2,76	0,4971	3,60	0,499841
2,16	0,4846	2,46	0,4931	2,78	0,4973	3,80	0,499928
2,18	0,4854	2,48	0,4934	2,80	0,4974	4,00	0,499968
2,20	0,4861	2,50	0,4938	2,82	0,4976	4,50	0,499997
		2,52	0,4941			5,00	0,500000

Додаток 3

Таблиця значений $t_\gamma = t(\gamma, n)$

n \ γ	0,95	0,99	0,999	n \ γ	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Додаток 4

Таблиця значений $q = q(\gamma, n)$

n \ γ	0,95	0,99	0,999	n \ γ	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,68	1,01	150	0,115	0,160	0,221
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Додаток 5

Таблица значений χ^2 в зависимости от p и k

k	0,05	0,01	0,001
1	3,84	6,64	10,83
2	5,99	9,21	13,82
3	7,82	11,34	16,27
4	9,49	13,28	18,46
5	11,07	15,09	20,5
6	12,59	16,81	22,5
7	14,07	18,48	24,3
8	15,51	20,1	26,1
9	16,92	21,7	27,9
10	18,31	23,2	29,6
11	19,68	24,7	31,3
12	21,0	26,2	32,9
13	22,4	27,7	34,6
14	23,7	29,1	36,1
15	25,0	30,6	37,7
16	26,3	32,0	39,3
17	27,6	33,4	40,8
18	28,9	34,8	42,3
19	30,1	36,2	43,8
20	31,4	37,6	45,3
21	32,7	38,9	46,8
22	33,9	40,3	48,3
23	35,2	41,6	49,7
24	36,4	43,0	51,2
25	37,7	44,3	52,6
26	38,9	45,6	54,1
27	40,1	47,0	55,5
28	41,3	48,3	56,9
29	42,6	49,6	58,3
30	43,8	50,9	59,7

Методичне видання

Кузьмич Л.В.

МАТЕМАТИЧНА СТАТИСТИКА

Методичні вказівки

ISBN 966-7403-50-5

Технічний редактор – Дудченко С.Г.

Здано до набору 4.07.00. Підписано до друку 10.07.00.
Формат 60x84 1/16. Папір офсетний. Друк різографія.
Гарнітура Arial. Умовн.друк.арк. 1,25. Наклад 300 прим.

Віддруковано у ТОВ "Айлант"
73000, Україна, м.Херсон, пров.Пугачова, 5/20.
Тел.: 26-67-22.