

VII. ДОДАТКИ ДО ТЕКСТІВ

5.3.N.1. СТАТИСТИЧНИЙ АНАЛІЗ РЕЗУЛЬТАТІВ ХІМІЧНОГО ЕКСПЕРИМЕНТУ

ЗМІСТ

1. ВИБІРКА

- 1.1. Середнє значення та стандартне відхилення
- 1.2. Перевірка однорідності вибірки. Виявлення значень варіант, що випадають
 - 1.2.1. Вибірки, малі за обсягом: критерій Діксона (Q-критерій)
 - 1.2.2. Вибірки, середні за обсягом: випробування Граббса
 - 1.2.3. Вибірки, великі за обсягом: 3s-критерій
 - 1.2.4. Загальні зауваження щодо виявлення варіант, що випадають
 - 1.2.4.1. Вибірки, малі за обсягом
 - 1.2.4.2. Вибірки, великі за обсягом
- 1.3. Об'єднання вибірок
 - 1.3.1. Об'єднана дисперсія й об'єднане середнє
 - 1.3.2. Критерій Бартлета
 - 1.3.3. Критерій Кокрена
 - ▼ 1.3.4. Критерій Левене
 - 1.3.4.1. Статистична модель
 - 1.3.4.2. Приклад: Перевірка однорідності дисперсій за критерієм Левене результатів титриметричного кількісного визначення кислоти аскорбінової
 - 1.3.5. Загальні зауваження щодо перевірки гіпотези однорідності дисперсій
 - 1.3.5.1. Проблема коректності і сумісності критеріїв Кокрена і Бартлета
 - 1.3.5.2. Перевірка гіпотези нормальності вибірок

1.3.5.3. Проблема малих вибірок з різним числом ступенів свободи▲

- 1.4. Довірчі інтервали й оцінювання їх величини
- 1.5. Однобічні та двобічні довірчі інтервали

2. МЕТРОЛОГІЧНІ ХАРАКТЕРИСТИКИ МЕТОДИКИ АНАЛІЗУ

- 2.1. Представлення метрологічних характеристик
- 2.2. Формулювання аналітичної задачі
- 2.3. Особливості контролю якості лікарських засобів за показником «Кількісне визначення»
 - 2.3.1. Загальні положення
 - 2.3.2. Доказовий підхід
 - 2.3.3. Підтверджувальний підхід
- 2.4. Оцінка значущості систематичної похибки
 - 2.4.1. Статистична значущість систематичної похибки
 - 2.4.2. Практична значущість систематичної похибки
 - 2.4.2.1. Принцип незначущості
 - 2.4.2.2. Критерій практичної незначущості систематичної похибки

3. ПОРІВНЯННЯ ДВОХ МЕТОДИК АНАЛІЗУ ЗА ВІДТВОРЮВАНІСТЮ

4. МЕТРОЛОГІЧНА ХАРАКТЕРИСТИКА СЕРЕДНЬОГО РЕЗУЛЬТАТУ

5. ПОРІВНЯННЯ СЕРЕДНІХ РЕЗУЛЬТАТІВ ДВОХ ВИБІРОК

- 5.1. Розходження дисперсій s_1^2 і s_2^2 статистично невірогідно
- 5.2. Розходження дисперсій s_1^2 і s_2^2 статистично вірогідно
- 5.3. Відоме точне значення величини A
- 5.4. Використання довірчих інтервалів

6. ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ АНАЛІЗУ

- 6.1. Оцінка збіжності результатів паралельних випробувань
- 6.2. Забезпечення якості результатів

- 6.2.1. Кваліфікація мірного посуду
 - 6.2.2. Придатність хроматографічної системи
 - 6.3. Гарантія якості продукції
 - 6.3.1. Метрологічно атестована методика
 - 6.3.2. Валідована методика
 - 6.3.3. Урахування факторів неоднорідності для дозованих одиниць
 - 6.4. Оцінка результатів тестування
 - 6.4.1. Розрахунок метрологічних характеристик міжлабораторної вибірки
 - 6.4.2. Оцінка якості результатів учасників тестування
 - 6.4.3. Внутрішньолабораторне тестування
 - 6.4.4. Зовнішнє тестування
 - 6.4.4.1. Використання метрологічних характеристик результатів вибірки учасників
 - 6.4.4.2. Використання максимально допустимої невизначеності методики аналізу
 - 6.4.4.2.1. Наявність загальної систематичної похибки результатів учасників
 - 6.4.4.2.2. Загальна характеристика якості результатів учасників
- 7. РОЗРАХУНОК І СТАТИСТИЧНА ОЦІНКА ПАРАМЕТРІВ ЛІНІЙНОЇ ЗАЛЕЖНОСТІ**
- 8. ПРИКЛАДИ**
- 8.1. Обчислення середнього значення та дисперсії
 - 8.2. Перевірка однорідності вибірки малого обсягу
 - 8.2.1. Q-критерій
 - 8.2.2. Критерій Граббса
 - 8.3. Обчислення довірчих інтервалів і невизначеностей вимірювань
 - 8.4. Перевірка гіпотези рівності дисперсій
 - 8.4.1. Об'єднання результатів вибірок, різних за обсягом
 - 8.4.2. Об'єднання результатів вибірок, однакових за обсягом
- 8.5. Порівняння двох методик аналізу за прецизійністю і правильністю результатів. Статистична і практична значущість систематичної похибки**
- 8.5.1. Збіжність та наявність статистично значущої систематичної похибки
 - 8.5.2. Наявність практично значущої систематичної похибки
- 8.6. Порівняння середніх результатів двох вибірок**
- 8.6.1. Використання дисперсій
 - 8.6.2. Використання максимально допустимої невизначеності
- 8.7. Забезпечення якості результатів**
- 8.7.1. Кваліфікація градуйованої піпетки місткістю 5 мл
- 8.8. Гарантувальні допуски**
- 8.8.1. Метрологічно атестована методика
 - 8.8.2. Валідована методика
 - 8.8.3. Урахування факторів неоднорідності для дозованих одиниць
 - 8.8.4. Граничні допуски вмісту за специфікацією для дозованих одиниць
- 8.9. Оцінка результатів тестування**
- 8.9.1. Внутрішньолабораторне тестування персоналу за результатами верифікації градуйованої піпетки місткістю 5 мл
 - 8.9.2. Зовнішнє тестування: визначення вмісту домішки В лінкоміцину в субстанції лінкоміцину методом ВЕРХ
- 9. РОЗРАХУНОК НЕВИЗНАЧЕНОСТІ ФУНКЦІЇ ДЕКІЛЬКОХ ВИПАДКОВИХ ЗМІННИХ**
- 9.1. Лінійна модель
 - 9.1.1. Зважене середнє
 - 9.2. Підхід Велча — Саттертвейта
 - 9.3. Приклади розрахунків невизначеності функції декількох змінних
 - 9.3.1. Розрахунок невизначеності методики аналізу готового лікарського засобу за допомогою високоефективної рідинної хроматографії (ВЕРХ)
 - 9.3.1.1. Кінцева аналітична операція

9.3.1.2. Сумарна невизначеність пробопідготовки $\Delta_{SP,r}$

9.3.1.3. Розрахунок сумарної невизначеності методики аналізу $\Delta_{As,r}$

9.3.2. Прогноз невизначеності спектрофотометричного аналізу готового лікарського засобу

9.3.3. Розрахунок середнього значення декількох нерівноточних вибірок

9.3.4. Порівняння застосування лінійної моделі і підходу Велча — Саттертвейта в граничному випадку

10. ДОДАТОК

Таблиця 10.1.1. Критичні значення однобічного контрольного критерію Діксона $Q(n, \alpha)$

Таблиця 10.1.2. Числові значення критерію Граббса (ESD_{crit})

Таблиця 10.2. Числові значення критерію Стьюдента $t(p, \nu)$

Таблиця 10.3. Відсоткові точки розподілу $\chi^2(P_1, \nu)$

Таблиця 10.4. Критерій Кокрена

Таблиця 10.5. Відсоткові точки розподілу Фішера ($F(P_1, \nu_1, \nu_2)$ — розподіл)

Таблиця 10.6. Відсоткові точки вибіркового коефіцієнта кореляції r

11. ДЕЯКІ ДИСКРЕТНІ СТАТИСТИЧНІ РОЗПОДІЛИ

11.1. Біноміальний розподіл

11.1.1. Визначення

11.1.2. Застосування у фармацевтичному аналізі

11.1.2.1. Гранично допустима кількість дефектних одиниць під час підтвердження якості

11.1.2.2. Довірчі межі для параметра p біноміального розподілу

11.2. Розподіл Пуассона

Приклад 11.1. Підтверджувальний вибірковий контроль якості ін'єкцій за механічними вклученнями

Приклад 11.2. Узагальнене оцінювання якості результатів учасників міжлабораторного тестування

12. ЗАСТОСУВАННЯ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ У РЕГРЕСІЙНОМУ АНАЛІЗІ

12.1. Загальне формулювання методу найменших квадратів

12.2. Проблема нерівноточності значень ординат

12.2.1. Допущення рівноточності ординат

12.2.2. Загальний випадок

12.2.3. Допущення рівності відносних стандартних відхилень ординат

12.3. Пряме числове розв'язання співвідношень (12.3) або (12.7)

12.4. Використання часткових похідних регресійної функції за параметрами

12.5. Лінійний метод найменших квадратів (ЛМНК)

12.6. Лінійний зважений метод найменших квадратів (ЛЗМНК)

12.7. Нелінійний метод найменших квадратів (НМНК)

12.8. Нелінійний зважений метод найменших квадратів (НЗМНК)

12.9. Приклади

12.9.1. Валідація методики кількісного визначення таблеток амброксолу: розрахунок лінійної залежності методом найменших квадратів

12.9.2. Дослідження значимості впливу швидкості і сили пресування на однорідність маси і однорідність дозування одиниць ядер таблеток дезлоратадину

12.9.3. Лінійний зважений метод найменших квадратів: побудова калібрувальної прямої в припущенні рівності відносних невизначеностей ординат

12.9.4. Застосування нелінійного методу найменших квадратів для опису кінетики розчинення зразків субстанції фенсукциналау з різним розміром частинок

13. ЗАСТОСУВАННЯ ДИСПЕРСІЙНОГО АНАЛІЗУ ДЛЯ ІНТЕРПРЕТАЦІЇ АНАЛІТИЧНИХ ДАНИХ

13.1. Порівняння середніх значень двох вибірок: t -критерій

13.2. Порівняння середніх значень трьох і більше вибірок: дисперсійний аналіз

13.2.1. Порівняння дисперсійного і регресійного аналізу	13.10.2.2. Парні порівняння
13.3. Загальні відомості про дисперсійний аналіз	13.10.2.3. Контрасти
13.4. Вихідні положення дисперсійного аналізу	13.10.2.3.1. Контрасти комбінацій умов
13.5. Статистична модель дисперсійного аналізу	13.10.2.3.2. Контрасти сум
13.5.1. Однофакторний дисперсійний аналіз	13.10.2.3.3. Контрасти користувача
13.5.2. Багатофакторний дисперсійний аналіз	13.10.3. Відпрацювання технології отримання таблеток: взаємозв'язок між однорідністю маси і дозування
13.6. Коефіцієнт детермінації і загальний індекс кореляції	13.10.3.1. Дисперсійний аналіз однорідності маси
13.7. Загальна схема дисперсійного аналізу	13.10.3.2. Дисперсійний аналіз однорідності вмісту
13.8. Апостеріорний аналіз результатів: парні порівняння	13.10.3.3. Дисперсійний аналіз величини $100 \times UC_{ij} / UM_{ij} \%$
13.8.1. Постановка завдання	13.10.3.4. Регресійний аналіз взаємозв'язку між однорідністю вмісту і однорідністю маси
13.8.2. Проблеми множинних парних порівнянь	13.10.3.4.1. Вихідні положення
13.8.2.1. Проблема сумісності парних і групових критеріїв	13.10.3.4.2. Залежність кількісного вмісту від середньої маси таблеток
13.8.2.2. Проблема зростання ймовірності великих різниць з ростом кількості (k) вибірок	13.10.3.4.3. Обговорення результатів регресійного аналізу▲
13.8.3. Тест Тьюки	
13.8.4. Поправки Бонфероні і Холма-Бонфероні	
13.8.5. Процедура апостеріорного аналізу результатів методом парних порівнянь, що рекомендується	
13.9. Метод контрастів	
13.9.1. Визначення	
13.9.2. Контрасти комбінацій умов	
13.9.3. Контрасти сум	
13.9.4. Контрасти користувача	
13.9.5. Контрасти Гелмерта	
13.10. Приклади	
13.10.1. Порівняння середніх результатів двох вибірок, отриманих при кількісному визначенні субстанції кислоти аскорбінової	
13.10.2. Кількісне визначення кислоти аскорбінової методом титрування: інтерпретація результатів	
13.10.2.1. Дисперсійний аналіз результатів	

ПРИЙНЯТІ ПОЗНАЧЕННЯ

У цій статті для переважного використання прийняті такі позначення:

A	вимірювана величина;
a	вільний член лінійної залежності;
b	кутовий коефіцієнт лінійної залежності;
F	критерій Фішера;
$f(x, \mu, \sigma)$	функція щільності ймовірності нормального розподілу;
H_0	нульова гіпотеза;
H_1	альтернативна гіпотеза;
i	порядковий номер варіанти;
L	фактор, використовуваний під час оцінки збіжності результатів рівнобіжних визначень;
m, n	об'єми вибірки;

p	ймовірність;	X_i, Y_i	значення змінних x та y , обчислені з рівняння лінійної залежності;
P	довірча ймовірність без конкретизації постановки завдання, у відсотках;	\bar{x}, \bar{y}	середні значення вибірки (координати центра лінійної залежності);
P_2, P_1	довірча ймовірність відповідно в разі дво- й однобічного завдання, у відсотках;	x_i, y_i	i -та варіанта (i -та пара експериментальних значень x та y);
Q_1, Q_n	контрольні критерії для ідентифікації грубих похибок;	$\bar{x} \pm \Delta_{\bar{x}}$	граничні значення довірчого інтервалу середнього результату;
R	розмах варіювання;		
R_c	загальний індекс кореляції;	$x_i \pm \Delta_x$	граничні значення довірчого інтервалу результату одиничного визначення;
r	(лінійний) коефіцієнт кореляції;	Δ_x	напівширина довірчого інтервалу одиничного визначення;
$RSD = s_r \times 100 \%$	відносне стандартне відхилення, у відсотках;	$\Delta_{\bar{x}}$	напівширина довірчого інтервалу середнього результату;
$RSD_{\bar{x}} = s_{\bar{x},r} \times 100 \%$	відносне стандартне відхилення середнього результату, у відсотках;	$\Delta_{x,r}$	напівширина відносного довірчого інтервалу одиничного визначення;
RSD_p	об'єднане відносне стандартне відхилення, у відсотках;	$\Delta_{\bar{x},r}$	напівширина відносного довірчого інтервалу середнього результату;
s	стандартне відхилення;	$\Delta_{As,r}$	сумарна невизначеність аналізу;
s_r	відносне (щодо середнього результату) стандартне відхилення;	$\Delta_{FAO,r}$	невизначеність кінцевої аналітичної операції;
s^2	дисперсія;	$\Delta_{RS,r}$	невизначеність атестації стандартного зразка;
s_r^2	відносна дисперсія;	$\Delta_{SP,r}$	невизначеність пробопідготовки;
$s_{\bar{x}}$	стандартне відхилення середнього результату;	δ	відносна величина систематичної похибки;
$s_{\bar{x},r}$	відносне (щодо середнього результату) стандартне відхилення середнього результату;	$\varepsilon, \bar{\varepsilon}$	відносні невизначеності відповідно результату окремого визначення і середнього результату;
s_{lg}	логарифмічне стандартне відхилення;	μ	справжнє значення вимірюваної величини;
s_{lg}^2	логарифмічна дисперсія;	ν	число ступенів свободи; перемінний обсяг вибірки в разі послідовного аналізу;
$s_{lg\bar{x}}$	логарифмічне стандартне відхилення середнього результату;	ν_p	об'єднане число ступенів свободи;
s_p	об'єднане стандартне відхилення;	ν_{eff}	«ефективне» число ступенів свободи в підході Велча — Саттертвейта;
s_0^2, s_b^2, s_a^2	загальна дисперсія та дисперсії коефіцієнтів лінійної залежності;	Σ	знак підсумовування (сума);
▼ SD	стандартне відхилення при проведенні валідаційних досліджень, регресійного і дисперсійного аналізу; ▲	σ	стандартне відхилення генеральної сукупності,
t	критерій Стюдента;		
x, y	поточні координати в рівнянні лінійної залежності;		

σ^2 дисперсія генеральної сукупності;
 χ^2 критерій «хі-квадрат».

$$F(P, k - 1, N - k) = W = \frac{N - k}{k - 1} \times \frac{\sum_{i=1}^{i=k} n_i \times (Z_{i.} - Z_{..})^2}{\sum_{i=1}^{i=k} \sum_{j=1}^{n_i} (Z_{ij} - Z_{i.})^2}$$

1.3.4. Критерій Левене

Критерії Бартлета (1.3.2) і Кокрена (1.3.3) ґрунтуються на використанні тільки групових дисперсій і ступенів свободи. Це має свої переваги, але й недоліки.

Критерій Левене (Levene's test) був запропонований у 1960 році як альтернатива критерію Бартлета (див. розділ 1.3.2) для перевірки гіпотези рівності (однорідності) дисперсій. На відміну від критеріїв Бартлета і Кокрена, він використовує весь масив даних, а не тільки дисперсії і числа ступенів свободи. Це робить його більш представницький і менш чутливим до відхилень від нормального розподілу вибірок, що досліджуються. Тому він часто застосовується при порівнянні вибірових середніх (див. розділ 5).

Особливо доцільно застосування критерію Левене в дисперсійному аналізі (див. розділ 13). Це пов'язано з низкою факторів:

1. Порівняння групових (вибіркових) середніх у дисперсійному аналізі ґрунтується на гіпотезі однорідності (рівності) внутрішньо групових дисперсій.
2. Критерій Левене (див. нижче), як і дисперсійний аналіз (див. розділ 13), ґрунтується на прямому використанні критерію Фішера, тому отримані за ним результати є порівнянними. Це вигідно відрізняє його від критеріїв Кокрена і Бартлета.
3. Критерій Левене легко реалізується в *Excel*, що забезпечує його просліджуваність.

Недоліки критерію Левене:

1. Критерій Левене використовує весь масив даних, що не завжди можливо чи зручно. Наприклад, при проведенні міжлабораторного тестування часто надаються тільки стандартні відхилення і середні, що робить неможливим застосування критерію Левене.
2. Якщо кількість груп є великим (наприклад, при міжлабораторному тестуванні вона часто перевищує 50), то застосування критерію Левене стає досить громіздким. Застосування критеріїв Бартлета (в загальному випадку) і Кокрена (при рівному числі ступенів свободи) має у цьому випадку переваги.

1.3.4.1. Статистична модель

Визначення:

Нульова гіпотеза: $H_0: \sigma_1 = \sigma_2 = \dots = \sigma_k$.

Альтернативна гіпотеза: $H_1: \sigma_i \neq \sigma_j$ хоча б для однієї пари (i, j) .

Критеріальна статистика Левене визначається таким чином:

	$P(F = W) < P_{crit}$	— індикс вибірки (групи), $1 \leq i \leq k$.
де i		— кількість вибірок (груп).
k		— індикс випадкової змінної усередині групи, $1 \leq j \leq n_i$.
j		— обсяг i -ої вибірки (групи).
n_i		— загальний обсяг усіх вибірок (груп).
$N = \sum n_i$		— j -оє значення випадкової змінної в i -ій вибірці (групі).
Y_{ij}		— середнє значення величини Y в i -ій вибірці (групі).
$Y_{i.}$		— середнє значення величини Y за всіма k вибірками (групами).
$Y_{..}$		— абсолютне значення (модуль) відхилення Y_{ij} від середнього значення $Y_{i.}$ в i -ій вибірці (групі).
$Z_{ij} = Y_{ij} - Y_{i.} $		— середнє значення величини Z_{ij} в i -ій вибірці (групі).
$Z_{i.}$		— середнє значення величини Z_{ij} за усіма вибірками (групами).
$Z_{..}$		— значення критерію Фішера для ступенів свободи $k - 1$ і $N - k$ і ймовірності (значущості) P .
$F(P, k - 1, N - k)$		— ймовірність, яка відповідає критерію Фішера $F(P, k - 1, N - k) = W$. Часто використовується також позначення a .
$P(F=W, k - 1, N - k)$		— значення ймовірності (значущості), яке прийняте за критичне.
P_{crit}		

Нульова гіпотеза рівності дисперсій відкидається, якщо ймовірність $P(F = W)$ менше критичної величини P_{crit} . Відмінність дисперсій вважається значущою, якщо $P(F) < P_{crit} = 5\%$, і високозначущою, якщо $P(F=W, k - 1, N - k) < P_{crit} = 1\%$.

Можна зразу розрахувати критичне значення $F(P_{crit}, k - 1, N - k)$ для даного значення P_{crit} і використовувати його як W_{crit} . Однак, приймаючи до уваги широке розповсюдження вбудованих статистичних програм (наприклад, в *Excel*), значно зручніше і ін-

формативніше використовувати розрахунок величини $P(F = W, k - 1, N - k)$.

1.3.4.2. Приклад: Перевірка однорідності дисперсій за критерієм Левене результатів титриметричного кількісного визначення кислоти аскорбінової

Кількісне визначення субстанції кислоти аскорбінової проводилось методом йодометричного титрування. Для рандомізації досліджень, при аналізі тестових зразків використовували різні реактиви і бюретки. Результати титриметричного аналізу в 5 серіях (групах) дослідів ($k = 5$) представлені в Табл. 13.П-1.

Для оцінки гіпотези рівності середніх результатів в 5 різних групах застосували дисперсійний аналіз (див. розділ 13). Оскільки дисперсійний аналіз ґрунтується на гіпотезі однорідності дисперсій, то для її перевірки застосували критерій Левене. Результати такої перевірки даних Табл. 13.П-1 представлені в Табл. 13.П-3.

Як видно, дисперсії результатів різних дослідів (груп) витримують випробування Левене на однорідність.

1.3.5. Загальні зауваження щодо перевірки гіпотези однорідності дисперсій

Перевірка однорідності дисперсій на практиці викликає певні труднощі, які є специфічними для кожного критерію.

1.3.5.1. Проблема коректності і сумісності критеріїв Кокрена і Бартлета

Критерій Кокрена (див. розділ 1.3.3) є найпростішим тестом на перевірку гіпотези однорідності дисперсій і тому широко застосовується при рівності ступенів свободи для досліджуваних дисперсій. В цих же умовах можна застосовувати і більш загальний, але більш складний критерій Бартлета (див. розділ 1.3.2). Ці критерії ґрунтуються на різних принципах і тому можуть призводити до протилежних висновків. Який висновок щодо однорідності дисперсій є коректним у цьому випадку? Розглянемо модельний приклад.

ПРИКЛАД 1.1

Виберемо явно неоднорідну сукупність $g = 6$ стандартних відхилень: $s_i = 0.01, 0.01, 0.01, 0.65, 0.95$ і 1.00 з однаковим числом ступенів свободи ν_k . Чи виконуються для цієї сукупності критерії однорідності дисперсій Кокрена і Бартлета?

Як видно (Табл. 1.1), критерій Бартлета чітко вказує, що ця сукупність стандартних відхилень не витримує вимоги на однорідність дисперсій для будь-якого числа ступенів свободи. Одночасно, критерій Кокрена дає абсурдний висновок про однорідність цієї сукупності дисперсій для числа ступенів свободи $\nu < 6$. Це пов'язано з тим, що критерій Кокрена не є симетричним по відношенню до відхилення дисперсій – він є нечутливим до відхилень дисперсій в меншу сторону. На відміну від критерію Кокрена, критерій Бартлета є симетричним по відношенню до відхилень дисперсій в меншу і більшу сторони, що робить його застосування більш надійним у загальному випадку.

Відзначимо, що цей висновок не залежить від того, відповідають чи ні вихідні вибірки вимогам нормальності розподілу.

В цьому «гострому» випадку некоректність застосування критерію Кокрена є очевидною. Однак в загальному випадку такої очевидності немає і незрозуміло, який висновок щодо однорідності дисперсій треба робити у випадку розбіжності між критеріями Кокрена і Бартлета.

1.3.5.2. Перевірки гіпотези нормальності вибірок

Критерії Кокрена і Бартлета є чутливими до відхилень вибірок від нормальності, тому перед їх застосуванням гіпотезу нормальності слід перевірити. На практиці часто доводиться мати справу з великою кількістю вибірок (у випадку міжлабораторного тестування звичайним є число вибірок більше 50), які мають невелике число ступенів свободи ($\nu = 2-4$). Процедура перевірки в цьому випадку гіпотези нормальності для всіх вибірок є дуже громіздкою і ненадійною.

Таблиця 1.1

Порівняння застосування критеріїв Кокрена і Бартлета для різних ступенів свободи ν_k .

	$\nu_k = 1$	$\nu_k = 2$	$\nu_k = 3$	$\nu_k = 4$	$\nu_k = 5$	$\nu_k = 6$
Критерій Кокрена ($G < \max G$)						
G	0.430					
$\max G$	0.781	0.616	0.532	0.480	0.445	0.418
Критерій Бартлета ($\chi^2 < \max \chi^2$)*						
χ^2	22.91	45.82	68.74	91.65	114.56	137.47
$\max \chi^2$	3.84	5.99	7.82	9.49	11.07	12.59

* Критерій Бартлета є застосовним при $\nu_k > 3$, тому його значення для $\nu_k \leq 3$ наведені як інформаційні

1.3.5.3. Проблема малих вибірок з різним числом ступенів свободи

На практиці (наприклад, при проведенні міжлабораторного тестування) часто доводиться мати справу з великою кількістю вибірок, серед яких є вибірки з невеликим і різним числом ступенів свободи ($\nu \leq 3$). Критерій Кокрена в цьому випадку не застосовний, оскільки число ступенів свободи не є однаковим. Не застосовний і критерій Бартлета, оскільки число ступенів свободи $\nu \leq 3$. Як перевіряти у цьому випадку гіпотезу однорідності дисперсій? Певно, за відсутністю кращого, можна застосувати критерій Бартлета. Однак, наскільки коректним у цьому випадку є висновок щодо однорідності дисперсій?

1.3.5.3. Проблема великого числа вибірок для критерію Левене

Цей критерій не має недоліків критеріїв Бартлета (вимоги до мінімального числа ступенів свободи $\nu \leq 3$) і Кокрена (вимоги до однаковості числа ступенів свободи для всіх дисперсій, а також несиметричність до відхилень дисперсій в меншу і більшу сторони). Однак застосування критерію Левене для великого числа вибірок є дуже громіздким.

Так, із Табл. 13.П-3 видно, що для 5 груп критерій Левене використовує 17 стовбців. У випадку 50 груп число стовбців зростає до 152, що позбавляє цей тест просліджуваності.

13. ЗАСТОСУВАННЯ ДИСПЕРСІЙНОГО АНАЛІЗУ ДЛЯ ІНТЕРПРЕТАЦІЇ АНАЛІТИЧНИХ ДАНИХ

13.1. ПОРІВНЯННЯ СЕРЕДНІХ ЗНАЧЕНЬ ДВОХ ВИБІРОК: *t*-КРИТЕРІЙ

На практиці часто виникає завдання порівняння середніх значень декількох вибірок – чи можна їх вважати вибірками із однієї генеральної сукупності?

Статистичні підходи для порівняння середніх значень двох вибірок розглянуті в розділі 5.3 глави 5.3.N.1. Оцінку значущості різниці середніх проводять за допомогою *t*-критерію, приймаючи за нульову гіпотезу рівність двох вибірових середніх.

Основна ідея *t*-критерію – різниця вибірових (групових) середніх не має перевищувати довірчий інтервал середнього значення, розрахований на основі об'єднаного стандартного відхилення середнього двох вибірок. Якщо ця різниця вище, то різницю середніх двох вибірок вважають статистично значущою на вибраному рівні значущості α (ймовірність того, що середні двох вибірок належать одній і тій же генеральній сукупності), і нульову гіпотезу (рівність двох вибірових середніх) відкидають. Різницю на рівні $\alpha < 0.05$ (або $< 5.0\%$) вважають значущою, а на рівні $\alpha < 0.01$ (або $< 1.0\%$) – високозначущою.

При застосуванні *t*-критерію розраховують різницю вибірових середніх, об'єднане стандартне відхилення середнього (s_p) і об'єднане число ступенів свободи ν' . На основі цих величин розраховують *t*-критерій, який порівнюють з табличним значенням для заданої ймовірності. Якщо фактична величина *t* перевищує табличне критичне значення для вибраного рівня значущості α , то нульову гіпотезу рівності вибірових (групових) середніх відкидають. Цей підхід застосовний як у випадку статистичної незначущості різниці дисперсій двох вибірок (розділ 5.1 глави 5.3.N.1), так і за статистичної незначущості такої різниці (розділ 5.2 глави 5.3.N.1).

Розрахунок звичайно проводять з використанням комп'ютерних програм (наприклад, в *Excel*). Тому доцільно відразу розраховувати ймовірність $p(>t)$ (або $P(>t)\%$) того, що обидва середніх належать до однієї генеральної сукупності, і порівнюють її з вибраним рівнем значущості α . В цьому випадку краще відразу застосовувати більш загальний підхід розділу 5.2 (різниця вибірових дисперсій є значущою).

Типовий приклад застосування *t*-критерію для оцінки значущості різниці двох вибірових середніх у загальному випадку (різниця вибірових дисперсій є значущою) представлений в розділі 13.10.1.

13.2. ПОРІВНЯННЯ СЕРЕДНІХ ЗНАЧЕНЬ ТРЬОХ І БІЛЬШЕ ВИБІРОК: ДИСПЕРСІЙНИЙ АНАЛІЗ

Застосування *t*-критерію є ефективним підходом при порівнянні середніх результатів двох вибірок, але він не застосовний у тому випадку, коли число вибірок три і більше. Для вирішення таких завдань використовується дисперсійний аналіз, який можна розглядати як узагальнення *t*-критерію для випадку будь-якого числа вибірок. Основна його ідея – оцінка значущості різниці між середніми результатами $k \geq 3$ вибірок (груп) проводиться на основі порівняння за критерієм Фішера дисперсії середніх результатів вибірок (груп) і об'єднаної внутрішньогрупової дисперсії.

Внутрішньогрупова (тобто всередині кожної групи) варіація результатів аналізу характеризує випадкову (або, як її часто звать в дисперсійному аналізі, «непояснену») невизначеність методики аналізу. Різниця між середніми результатами груп характеризує систематичну (для кожної групи) складову невизначеності методики аналізу.

На відміну від випадку двох вибірок, обов'язковою вимогою застосування дисперсійного аналізу є однорідність (статистична незначущість різниці) дисперсій вибірок.

Як видно, ідеологія дисперсійного аналізу шпунтується на *F*-критерії, що дещо відрізняється від ідеології *t*-критерію. В той же час, збільшення кількості вибірок розширює можливості дисперсійного аналізу.

лізу, який дозволяє отримати суттєво більшу інформацію, ніж t -критерій.

13.2.1. Порівняння дисперсійного і регресійного аналізу

В розділі 12 розглядається використання регресійного аналізу для математичного опису залежності експериментально отриманої залежної змінної («функція відгуку», «цільова функція» тощо) від незалежних змінних (факторів), значення яких визначаються умовами проведення експерименту. Вигляд математичної функції задається наперед (априорно). Очевидною вимогою застосування регресійного аналізу є *чисельне* визначення як залежної («функція відгуку»), так і незалежних змінних (факторів).

Кількість незалежних змінних (факторів), у загальному випадку, може бути довільним, а самі регресійні залежності можуть бути як лінійними, так і нелінійними. Так, в розділі 12 представлені приклади математичного опису таких регресій:

- однофакторна лінійна залежність оптичної густини від концентрації в нормалізованих координатах (розділ 12.9.1),
- двофакторна лінійна залежність відносної дисперсії однорідності дозованих одиниць від сили і швидкості пресування (розділ 12.9.2),
- двофакторна нелінійна експоненційна залежність розчинності від розміру частинок і часу розчинення (розділ 12.9.4).

Як видно, завдання регресійного аналізу — це математичний опис експериментальної залежності функції відгуку (залежної змінної) від незалежних змінних (факторів). В цьому випадку залежна і незалежна змінні задаються в чисельному вигляді. Такі регресійні залежності, особливо лінійні, широко застосовуються в аналітичній хімії і, зокрема, в фармацевтичному аналізі.

Однак на практиці доводиться стикатися з випадком, коли залежна змінна (функція відгуку) задається в *чисельному вигляді*, а незалежні змінні (фактори) — *якісно* (категоріально). В цьому випадку завданням є не математичний опис зв'язку залежної і незалежної змінних (як в регресійному аналізі), а оцінка статистичної значущості впливу різних факторів на функцію відгуку. Такі питання вирішує дисперсійний аналіз.

Типове завдання дисперсійного аналізу представлено в Табл. 13.П-1. В ній наведені результати (Y_{ij}) йодометричного кількісного визначення кислоти аскорбінової в субстанції. Аналіз проводили для $k = 5$ серій (груп) дослідів з використанням різного аналітичного посуду і реактивів. Кількість паралельних визначень (n_i) в групах варіювалась від 3 до 7. Необхідно з'ясувати, наскільки статистично значущими є розбіжності між групами.

В термінології дисперсійного аналізу, в цьому прикладі досліджується залежність однієї функції від-

гуку (результати кількісного визначення у відсотках до номінального значення — Y_{ij}) від одного якісного фактора, який має $k = 5$ рівнів (градацій, груп), що відповідають результатам аналізу $k = 5$ різних груп. В цьому випадку на кожному рівні i (тобто для кожної i -ої групи) проводять $n_i = 3 - 7$ паралельних визначень.

Якщо провести аналогію зі звичайною функцією однієї змінної, то k рівнів відповідають k чисельним значенням незалежної змінної. Для кожного i -ого значення незалежної змінної проводиться n_i паралельних вимірів (визначень). Наприклад, під час дослідження стабільності ми вивчаємо залежність кількісного визначення Y від часу для 5 точок ($t_i = 15, 30, 45, 60$ і 75 хвилин). Водночас для кожної i -ої часової точки проводимо $n_i = 3-7$ паралельних аналізів, отримуючи величини Y_{ij} . Як видно, різниця є лише в тому, що значення незалежної змінної задаються не чисельно, а категоріально.

У загальному випадку, якісних факторів може бути декілька. Зокрема, другим фактором може бути, наприклад, температура аналізу, яка задається на двох рівнях — 15°C і 25°C (допуски Фармакопеї).

Рішення таких завдань відноситься до дисперсійного аналізу (ANalysis Of VAriance — ANOVA), який може бути однофакторним (одна залежна кількісна змінна і одна незалежна якісна змінна), багатфакторним (одна залежна кількісна і декілька незалежних якісних змінних), а також багатовимірним (декілька залежних кількісних змінних). Відзначимо, що дисперсійний аналіз є однією з складових планування експерименту і в цій якості широко застосовується, наприклад, в біологічних випробуваннях і кількісних визначеннях (див. главу 5.3 Фармакопеї).

Приймаючи до уваги високі вимоги до невизначеності результатів і обмеженість обсягу вибірок, в аналітичній практиці звичайно обмежуються однофакторним дисперсійним аналізом (*One-way ANOVA*) (див. нижче).

Відзначимо, що дисперсійний і регресійний аналізи добре доповнюють один іншого (див. розділ 13.10.3).

13.3. ЗАГАЛЬНІ ВІДОМОСТІ ПРО ДИСПЕРСІЙНИЙ АНАЛІЗ

Концепція дисперсійного аналізу (*ANOVA*) була запропонована Рональдом Фішером (Ronald Fisher) в 1920 р.

ANOVA — це процедура порівняння середніх значень вибірок (в термінології Табл. 13.П-1 це величини M_j), на підставі якої можна зробити висновок про співвідношення середніх значень генеральних сукупностей, що їм відповідають. Найближчим і найпростішим аналогом *ANOVA* є t -критерій (критерій Стьюдента). Як вже говорилося вище, на відміну від

t -критерію, дисперсійний аналіз призначений для порівняння не двох, а трьох і більше вибірок.

Слово «дисперсійний» в назві вказує на те, що в процесі аналізу порівнюються компоненти дисперсії залежної змінної (в термінології Табл. 13.П-1 це величина Y_{ij}). Загальна варіація залежної змінної (Y) розкладається на дві складові: міжгрупова (факторна), яка обумовлена відмінністю груп (середніх значень M_i), і внутрішньогрупова (похибка), що обумовлена випадковими (неврахованими) причинами. Так, в Табл. 13.П-1 внутрішньогрупова варіація пов'язана зі збіжністю результатів всередині групи, а міжгрупова — з систематичною похибкою аналізу в різних групах. Чим більше відношення міжгрупової і внутрішньогрупової варіації (F -відношення – співвідношення (13.6)), тим більше різняться середні значення вибірок, що порівнюються, і тим вище статистична значущість цієї різниці.

В однофакторному дисперсійному аналізі (*One way ANOVA*) порівнюють між собою середні значення кожної вибірки і розраховують загальний рівень значущості різниці. Важливо відзначити, що висновок за результатами *ANOVA* торкається загальної відмінності всіх середніх, що порівнюються, без конкретизації того, які саме вибірки різняться, а які ні. Для ідентифікації пар вибірок, які значуще відрізняються одна від іншої середніми значеннями, використовують апостеріорні критерії парних порівнянь (*Post Hoc*), а для більш складних порівнянь — метод контрастів (*Contrasts*).

Таким чином, якщо регресійний аналіз оцінює значущість чисельного опису експериментальних даних за допомогою запропонованої математичної функцією (для цього використовують залишкове стандартне відхилення, коефіцієнт кореляції і стандартні відхилення параметрів кореляції – див. розділ 12 глави 5.3. *N. I*), то дисперсійний аналіз оцінює значущість впливу на експериментальні дані досліджуваних факторів, які задаються якісно.

13.4. ВИХІДНІ ПОЛОЖЕННЯ ДИСПЕРСІЙНОГО АНАЛІЗУ

Основною метою дисперсійного аналізу є дослідження значущості відмінності між груповими середніми значеннями. Для цього за F -критерієм порівнюють міжгрупову і об'єднану внутрішньогрупову дисперсії.

Вихідні допущення дисперсійного аналізу:

нормальний розподіл значень досліджуваної ознаки (в наведеному вище прикладі – результатів кількісного визначення) в генеральній сукупності;

випадковий і незалежний характер вибірки;

однорідність групових дисперсій, тобто генеральні внутрішньогрупові дисперсії одиничного результату всіх груп є однаковими:

$$SD_{wg,i}^2 = const, \quad i = 1 \dots k. \quad (13.1)$$

Гіпотезу однорідності дисперсій (13.1) можна перевірити за критеріями Бартлета (див. розділ 1.3.2), Кокрена (див. розділ 1.3.3) і Левене (див. розділ 1.3.4). Найширше застосовують в дисперсійному аналізі критерій Левене, оскільки він, також як і сам дисперсійний аналіз, заснований на використанні F -критерію. Це робить результати перевірки однорідності дисперсій і власне дисперсійного аналізу зіставними.

Нульовою гіпотезою в дисперсійному аналізі є твердження про рівність групових (вибіркових) середніх значень (M_i), тобто:

$$M_i = const, \quad i = 1 \dots k. \quad (13.2)$$

13.5. СТАТИСТИЧНА МОДЕЛЬ ДИСПЕРСІЙНОГО АНАЛІЗУ

Основне співвідношення дисперсійного аналізу можна отримати з таких тотожностей:

$$y_{ij} = y_{ij} = M + (M_i - M) + (y_{ij} - M_i).$$

$$(y_{ij} - M) = (y_{ij} - M_i) + (M_i - M).$$

Піднесемо обидві частини останнього рівняння в квадрат і проведемо підсумування по j (n_i – обсяг i -ої вибірки):

$$\begin{aligned} \sum_{j=1}^{j=n_i} (y_{ij} - M)^2 &= \sum_{j=1}^{j=n_i} (y_{ij} - M_i)^2 + \\ &+ \sum_{j=1}^{j=n_i} (M_i - M)^2 + (M_i - M) \times \sum_{j=1}^{j=n_i} (y_{ij} - M_i) = \\ &= \sum_{j=1}^{j=n_i} (y_{ij} - M_i)^2 + n_i \times (M_i - M)^2, \end{aligned}$$

оскільки $\sum (y_{ij} - M_i) = 0$.

Проводячи підсумування по i обох частин цього співвідношення, остаточно отримуємо:

$$\begin{aligned} \sum_{i,j=1}^{j=n_i, i=k} (y_{ij} - M)^2 &= SS_{tot} = \\ &= \sum_{i,j=1}^{j=n_i, i=k} (y_{ij} - M_i)^2 + \sum_{i=1}^{i=k} n_i \times (M_i - M)^2 = SS_{wg} + SS_{bg}. \end{aligned}$$

Як видно, загальну суму квадратів SS_{tot} можна розкласти на внутрішньогрупову суму квадратів SS_{wg} (її позначають також SS_{error}) і міжгрупову суму квадратів SS_{bg} (її позначають також як SS_{effect}):

$$SS_{tot} = SS_{bg} + SS_{wg} = SS_{effect} + SS_{error}. \quad (13.3)$$

Рівняння (13.3) є основним співвідношення дисперсійного аналізу.

13.5.1. Однофакторний дисперсійний аналіз

Ступені свободи для однофакторного дисперсійного аналізу знаходять із співвідношень (n_i – обсяг i -ої вибірки, N – сумарний обсяг всіх k вибірок):

$$N = \sum_{i=1}^{i=k} n_i.$$

$$v_{tot} = v_{bg} + v_{wg} = N - 1. \quad (13.4)$$

$$v_{wg} = N - k.$$

$$v_{bg} = v_{tot} - v_{wg} = k - 1.$$

З урахуванням співвідношень (13.3-13.4), внутрішньогрупову і міжгрупову дисперсії (в дисперсійному аналізі їх прийнято звати «середніми сумами квадратів» - MS) розраховують за співвідношеннями:

$$MS_{tot} = SS_{tot} / v_{tot}.$$

$$MS_{wg} = MS_{error} = SS_{wg} / v_{wg}.$$

$$MS_{bg} = MS_{effect} = SS_{bg} / v_{bg}. \quad (13.5)$$

$$SS_{bg} = SS_{tot} - SS_{wg} = \sum_{i=1}^{i=k} n_i \times (M_i - M)^2.$$

F -відношення і ймовірність, що йому відповідає, розраховують за рівняннями:

$$F_{bg/wg} = \frac{MS_{bg}}{MS_{wg}} = \frac{MS_{effect}}{MS_{error}} \leq F_{crit}(P, v_{bg}, v_{wg}). \quad (13.6)$$

$$P(F_{bg/wg}, v_{bg}, v_{wg}) \leq \alpha.$$

Величина $P(F_{bg/wg}, v_{bg}, v_{wg})$ є ймовірністю того, що всі вибірккові середні належать до однієї генеральної сукупності. Значення $P(F_{bg/wg}, v_{bg}, v_{wg}) < \alpha = 5.0\%$ говорить про значущу, а $P(F_{bg/wg}, v_{bg}, v_{wg}) < \alpha = 1.0\%$ — про високо значущу різницю групових середніх. Можна використовувати наперед розраховані величини F_{crit} для заданого рівня значущості (ймовірності) α і ступенів свободи v_{bg} і v_{wg} , але зручніше і інформативніше використовувати значення $P(F_{bg/wg}, v_{bg}, v_{wg})$.

13.5.2. Багатофакторний дисперсійний аналіз

На відміну від однофакторної моделі, де є одна міжгрупову суму квадратів, модель багатофакторного аналізу включає суми квадратів для кожного фактора окрема і суми квадратів для всіх взаємодій між ними. Проведення багатофакторного дисперсійного аналізу вимагає планування експерименту, що суттєво ускладнює його застосування.

Проілюструємо проблеми, які обмежують застосування багатофакторного дисперсійного аналізу в фармацевтичному аналізі.

Нехай в двохфакторній (AB) моделі є k груп за фактором A і f груп за фактором B . Відповідно, міжгрупову суму квадратів SS_{bg} розкладається на суму квадратів фактора A , суму квадратів фактора B і суму квадратів взаємодії факторів A і B . Загальне співвідношення (13.3) у цьому випадку набуває вигляду:

$$SS_{tot} = SS_{bg} + SS_{wg} = (SS_A + SS_B + SS_{AB}) + SS_{wg}. \quad (13.7)$$

Ступені свободи для двохфакторного дисперсійного аналізу розкладається аналогічним чином:

$$v_{tot} = v_{bg} + v_{wg} = (v_A + v_B + v_{AB}) + v_{wg} = N - 1.$$

$$v_{wg} = N - f \times k.$$

$$v_{wg} = v_{tot} - v_{wg} = v_A + v_B + v_{AB} = f \times k - 1. \quad (13.8)$$

$$v_A = f - 1.$$

$$v_B = k - 1.$$

$$v_{AB} = (f - 1) \times (k - 1).$$

Відповідні середні суми квадратів MS (тобто дисперсії) мають вигляд:

$$MS_{tot} = SS_{tot} / v_{tot}$$

$$MS_A = SS_A / v_A$$

$$MS_B = SS_B / v_B$$

$$MS_{AB} = SS_{AB} / v_{AB} \quad (13.9)$$

$$MS_{error} = MS_{wg} = SS_{wg} / v_{wg}$$

$$MS_{effect} = MS_{bg} = SS_{bg} / v_{bg} = (SS_{tot} - SS_{wg}) / v_{wg}$$

Значущість кожної складової міжгрупової середньої суми квадратів (дисперсій) можна оцінити за критерієм Фішера:

$$F_{bg/wg} = \frac{MS_{bg}}{MS_{wg}} = \frac{MS_{effect}}{MS_{error}} \leq F_{crit}(P, v_{bg}, v_{wg}).$$

$$F_{A/wg} = \frac{MS_A}{MS_{wg}} = \frac{MS_A}{MS_{error}} \leq F_{crit}(P, v_A, v_{wg}). \quad (13.10)$$

$$F_{B/wg} = \frac{MS_B}{MS_{wg}} = \frac{MS_B}{MS_{error}} \leq F_{crit}(P, v_B, v_{wg}).$$

$$F_{AB/wg} = \frac{MS_{AB}}{MS_{wg}} = \frac{MS_{AB}}{MS_{error}} \leq F_{crit}(P, v_{AB}, v_{wg}).$$

Відповідно, трьохфакторна (ABC) модель має урахувати не тільки парні взаємодії факторів AB , AC , BC , але й взаємодії всіх трьох факторів ABC . Вираз (13.7) набуває у цьому випадку вигляд:

$$SS_{tot} = (SS_A + SS_B + SS_C + SS_{AB} + SS_{BC} + SS_{AC} + SS_{ABC}) + SS_{wg}. \quad (13.11)$$

Наведені приклади досить наглядно ілюструють ускладнення дисперсійного аналізу з ростом кількості факторів, що, разом з необхідністю планування експерименту, а також збільшення вимог до обсягу групових вибірок, знижує цінність багатофакторного дисперсійного аналізу для фармацевтичного аналізу.

13.6. КОЕФІЦІЄНТ ДЕТЕРМІНАЦІЇ І ЗАГАЛЬНИЙ ІНДЕКС КОРЕЛЯЦІЇ

Для оцінки відмінності групових середніх, в дисперсійному аналізі використовують також коефіцієнт детермінації η^2 . Він показує, яка доля загальної варіації залежної змінної (у вигляді суми квадратів) пов'язана з міжгруповою варіацією, яка викликана цим фактором. У випадку однофакторного дисперсійного аналізу, вираз для η^2 має вигляд:

$$\eta^2 = 1 - \frac{SS_{wg}}{SS_{tot}} = 1 - \frac{SS_{error}}{SS_{tot}} = \frac{SS_{bg}}{SS_{tot}} = \frac{SS_{effect}}{SS_{tot}}. \quad (13.12)$$

У випадку багатофакторного аналізу, рівняння (13.12) легко узагальнюється для внеску кожного компоненту SS_{bg} (вираз в дужках рівнянь (13.7) і (13.11)) в загальну суму квадратів SS_{tot} . Зокрема, для двохфакторного аналізу вираз для коефіцієнта детермінації (13.12), з урахуванням (13.7), можна представити у вигляді:

$$\begin{aligned} \eta^2 &= \frac{SS_{bg}}{SS_{tot}} = \frac{SS_A}{SS_{tot}} + \frac{SS_B}{SS_{tot}} + \frac{SS_{AB}}{SS_{tot}} = \\ &= \eta_A^2 + \eta_B^2 + \eta_{AB}^2. \end{aligned} \quad (13.13)$$

Коефіцієнт детермінації (його звать також «кореляційним співвідношенням Пірсона») є мірою впливу досліджуваного фактора (у випадку Табл. 13.П-1 це різні групи дослідів) на варіацію залежної змінної у (в розглянутому прикладі це величини Y_{ij}).

У випадку однофакторного дисперсійного аналізу, $\eta^2 = 0$ свідчить про те, що цей фактор не впливає на y , і загальна варіація у пов'язана виключно з внутрішньогруповою варіацією (*error* в Табл. 13.П-2). Якщо $\eta^2 = 1$, то внутрішньогрупова варіація (*error*) відсутня, і загальна варіація у пов'язана виключно з міжгруповою варіацією (*effect* в Табл. 13.П-2).

Кореляційні співвідношення (коефіцієнти детермінації) використовувати простіше, ніж суми квадратів, оскільки вони нормовані на одиницю (або на 100 %) і не залежать від шкали вимірювань. Однак, на відміну від квадрату загального індексу кореляції (7.3а, 12.5) R_c^2 , коефіцієнт детермінації η^2 не може бути використаний для перевірки гіпотез, оскільки він не включає число ступенів свободи. Тому для нього не існує критичних значень для заданої ймовірності. Величина η^2 носить інформаційний характер.

Зокрема, в Табл. 13.П-1 коефіцієнт детермінації $\eta^2 = 0.650$ (або 65.0 %), що свідчить про суттєвий вплив міжгрупових відмінностей на результати кількісного визначення. Значущість цього впливу перевіряється за критерієм Фішера (13.6).

Із співвідношення (13.13) видно, що η^2 легко розкладається на суму вкладів різних компонентів SS_{bg} , що робить його зручним для опису їх впливу в дисперсійному аналізі конкретного об'єкта.

Неважно бачити, що коефіцієнт детермінації η^2 за формою співпадає з квадратом загального індексу кореляції R_c^2 (12.5), який широко застосовується в регресійному аналізі для перевірки опису експериментальних даних за допомогою запропонованої математичної моделі. В термінології дисперсійного аналізу його можна представити у вигляді:

$$\begin{aligned} R_c^2 &= 1 - \frac{SD_{wg}^2}{SD_{tot}^2} = 1 - \frac{1}{F_{tot/wg}} \leq R_{c,crit}^2(P, \nu_{wg}), \\ F_{tot/wg} &= \frac{1}{1 - R_c^2}. \end{aligned} \quad (13.14)$$

Із співвідношень (13.14) видно, що квадрат загального індексу кореляції R_c^2 є лише нормованою на одиницю (або на 100 %) формою F -критерію, тобто він не несе (у порівнянні з ним) нової інформації. Крім того, на відміну від коефіцієнту детермінації η^2 , квадрат загального індексу кореляції R_c^2 не має адитивності подібно до (13.13). Тому, на відміну від регресійного аналізу (де застосування коефіцієнта кореляції є обов'язковим – див. розділ 12), в дисперсійному аналізі коефіцієнт кореляції застосовується рідко.

13.7. ЗАГАЛЬНА СХЕМА ДИСПЕРСІЙНОГО АНАЛІЗУ

В загальній схемі дисперсійного аналізу використовують не вибіркові дисперсії, а відповідні їм суми квадратів (*Sum of Squares - SS*) відхилень. Для вибірки обсягу n_i вибіркова дисперсія розраховується як сума квадратів відхилень від вибіркового середнього, поділена на число ступенів свободи, що дорівнює $n_i - 1$. Таким чином, за фіксованого обсягу вибірки n , дисперсія є функцією суми квадратів відхилень.

Адитивність вибірових дисперсій витримується тільки за однакових обсягах вибірок або для генеральних величин. В інших випадках вона є лише приблизною. На відміну від от них, сума квадратів позбавлена цих недоліків і має адитивність. Тому використання замість дисперсій суми квадратів дозволяє отримати уніфіковану схему загального підходу дисперсійного аналізу.

Метою дисперсійного аналізу є перевірка статистичної значущості відмінностей між груповими середніми. Ця перевірка проводиться за допомогою розбиття загальної суми квадратів на складові, одна з яких обумовлена внутрішньогруповою (*within group - wg*) невизначеністю (в дисперсійному аналізі її часто звать *похибкою*), а друга обумовлена відмінністю середніх значень, тобто міжгруповою (*beyond group - bg*) невизначеністю (в дисперсійному аналізі її часто звать *ефектом*). Ці складові суми квадратів потім використовують для розрахунків відповідних дисперсій (які в дисперсійному аналізі звать *середніми сумами квадратів*) і оцінки за їх допомогою статистичної значущості відмінності між вибірковими середніми значеннями. Якщо ця відмінність є статистично значущою (на заданому рівні значущості α), нульова гіпотеза (13.2) відкидається і приймається альтернативна гіпотеза про наявність різниці між груповими середніми (тобто що вони не належать до однієї генеральної сукупності).

Середню суму квадратів (тобто дисперсію), пов'язану з внутрішньогруповою варіацією ($MS_{wg} = MS_{error}$),

звичайно звать залишковою складовою або *середнім квадратом похибки*. Це пов'язано з тим, що при проведенні експерименту вона звичайно не може бути передбачена або пояснена. З іншого боку, середню суму квадратів (тобто. дисперсію), пов'язану з міжгруповою варіацією ($MS_{bg} = MS_{effect}$), звичайно звать *середнім квадратом ефекту*. На відміну від MS_{error} , її можна пояснити відмінністю між груповими середніми значеннями.

Перевірка значущості в дисперсійному аналізі ґрунтується на порівнянні складової дисперсії, обумовленої міжгруповою варіацією (*середній квадрат ефекту*, або $MS_{bg} = MS_{effect}$) і складової дисперсії, обумовленої внутрішньогруповою варіацією (*середній квадрат похибки*, або $MS_{wg} = MS_{error}$). Якщо вірною є нульова гіпотеза (13.2) (рівність групових середніх), то їх відношення не має бути значуще більше одиниці. Таке порівняння проводять за допомогою F -критерію (13.6).

13.8. АПОСТЕРІОРНИЙ АНАЛІЗ РЕЗУЛЬТАТИВ: ПАРНІ ПОРІВНЯННЯ

Застосовуючи однофакторний дисперсійний аналіз (*ANOVA*), ми перевіряємо нульову гіпотезу, що всі групи (вибірки), що порівнюються, належать одній генеральній сукупності, і тому їх генеральні середні значення однакові, тобто

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k. \quad (13.15)$$

Якщо нульову гіпотезу (13.15) не вдається відкинути на заданому рівні значущості (наприклад, $\alpha = 0.05$), то в подальшому апостеріорному аналізі немає необхідності. Але що робити, якщо нульова гіпотеза (13.15) відкидається? В цьому випадку ми робимо висновок, що середні значення досліджуваних груп значуще різняться (тобто досліджуваний фактор суттєво впливає на змінну, що нас цікавить). Це *єдиний висновок*, який можна зробити на підставі власне дисперсійного аналізу.

Однак звичайно цікаво з'ясувати, які саме групи (вибірки) обумовлюють ці розбіжності. Тобто, які вибірки належать до генеральної сукупності, а які ні. Щоб з'ясувати це, необхідно провести парні порівняння середніх значень досліджуваних груп.

Для цього використовують апостеріорні (*Post Hoc*) критерії порівняння пар вибірок. Термін «апостеріорні» означає, що цю процедуру проводять тільки після встановлення статистично достовірного результату однофакторного дисперсійного аналізу – тобто що міжгрупова варіація значуще вище внутрішньогрупової. Якщо відмінність результатів за *ANOVA* виявилася статистично незначущою, то застосування процедури парних порівнянь є некоректним.

13.8.1. Постановка завдання

Завдання ідентифікації пар вибірок (груп), які значуще (на заданому рівні значущості α) відрізняються одна від іншої середніми значеннями, є неоднозначною і подібна до завдання перевірки однорідності вибірок і виключення варіант, що випадають (див. розділ 1.2 глави 5.3. *N.I*). Однак ця подібність є зовнішньою. При апостеріорному аналізі не ставиться завдання виключення групових середніх, що випадають, з метою отримання більш точних значень загального середнього і його дисперсії. Адже за результатами проведеного дисперсійного аналізу вже доказано, що середні значення груп не належать до однієї генеральної сукупності. Тому виключення будь-яких групових середніх є некоректним, оскільки вони всі рівноцінні.

Завдання апостеріорного аналізу методом парних порівнянь – це ідентифікація тих пар вибірок, середні значення яких значуще відрізняються одна від іншої, тобто, не належать до однієї генеральної сукупності.

Оцінка значущості різниці між кожними двома груповими середніми проводиться за t -критерієм (див. розділ 5 глави 5.3. *N.I*). В цьому випадку для кожної пари (ij) групових середніх ми розраховуємо t_{ij} -значення і відповідну йому ймовірність $p_{ij}(>t)$ того, що дана пара (ij) належить до однієї генеральної сукупності. Для вибраного рівня значущості (α) різниця між парою групових середніх (ij) є значущою, якщо:

$$p_{ij}(>t) < \alpha. \quad (13.16)$$

Значуща різниця відповідає $\alpha = 0.05$ (5.0 %), а високозначуща – $\alpha = 0.01$ (1.0 %).

Пари, які відповідають даному критерію, не можна вважати вибірками з однієї генеральної сукупності на рівні значущості α .

Проводячи ранжування ймовірності p_{ij} для всіх пар середніх значень, можна знайти те найменше значення p_{ij} , яке ще не відповідає критерію (13.16). Такий підхід звать тестом на найменшу значущу різницю (*LSD* - Least significant difference test). Знаходження найменшої p_{ij} в більшості випадків і є завданням апостеріорного аналізу результатів методом парних порівнянь. Пари вибірок з меншою ймовірністю вже не можна вважати вибірками з однієї генеральної сукупності.

13.8.2. Проблеми множинних парних порівнянь

Значущість відмінності між середніми значеннями різних груп можна оцінити за допомогою t -критерію (див. розділ 5 глави 5.3. *N.I*). Цей підхід запропонував ще Фішер. У загальному випадку внутрішньогрупові дисперсії і обсяги вибірок різняться, тому доцільно відразу застосовувати співвідношення

(5.7-5.9) розділу 5.2. Однак при цьому виникають такі перешкоди:

13.8.2.1. Проблема сумісності парних і групових критеріїв

Дисперсії вибірок можуть відповідати критеріям міжгрупової однорідності за Кокреном, Бартлетом або Левене (ця вимога є обов'язковою для застосування ANOVA), але не витримувати випробування за Фішером на рівність дисперсій конкретних пар вибірок. Це пов'язано з тим, що малі і великі дисперсії нерівноправні у вищевказаних групових критеріях, значення яких визначаються, здебільшого, великими дисперсіями. Це наочно видно на прикладі критерію Кокрена (1.19), на значення якого мало впливають малі дисперсії.

У випадку ж парних порівнянь за Фішером великі і малі дисперсії повністю рівноправні. Це призводить до великих значень відношень Фішера і, відповідно, значущим відмінностям дисперсій.

Так, в Табл. 13.П-1 є два невеликих вибірових стандартних відхилень ($SD_3 = 0.026$ и $SD_5 = 0.052$), які не впливають на відповідність дисперсій всіх вибірок критеріям Бартлета і Левене. Однак саме вони призводять до того, що у випадку парних порівнянь 6 з 10 пар не відповідають критерію Фішера на рівність дисперсій на рівні значущості $\alpha = 5.0 \%$ (див. Табл. 13.П-4).

Проблему сумісності парних і групових критеріїв найбільш природно вирішує тест Тьюки (*Tukey's honestly significant difference test*, або просто *Tukey's HSD test*).

13.8.2.2. Проблема росту ймовірності великих відмінностей з ростом кількості (k) вибірок

Це питання вже обговорювалося в розділі 1.2.3 глави 5.3.N.1. Якщо ми маємо k вибірок, то кількість комбінацій пар вибірок дорівнює числу сполучень із k по 2, тобто $m = k \times (k - 1) / 2$. Зокрема, в Табл. 13.П-1 маємо $k = 5$, і число комбінацій дорівнює $m = 5 \times (5-1)/2 = 10$. Для рівня значущості кожного порівняння $\alpha = 0.05$ (типового для аналітичної практики), існує ймовірність $1 - (1 - 0.05)^{10} = 0.40$ (тобто, 40 %) того, що хоча б 1 із 10 тестів покаже значущий результат випадково. З ростом числа (k) вибірок така ймовірність стрімко зростає. Відповідно зростає і похибка першого роду – ймовірність хибно позитивних результатів (тобто хибного визнання значущим відхилення, яке насправді таким не є).

Ця проблема має загальний характер. Для її вирішення запропоновані різні підходи, серед яких можна виділити поправки Бонфероні (Bonferroni) і Холма-Бонфероні (Holm-Bonferroni), які розглядаються нижче.

13.8.3. Тест Тьюки

Критерій Тьюки використовують для перевірки нульової гіпотези (13.15) проти альтернативної гіпотези

$$H_1: \mu_i \neq \mu_j, \quad i, j = 1 \dots k. \quad (13.17)$$

Тут індекси i і j означають будь-які дві групи, що порівнюються. При наявності k груп необхідно виконати $m = k \times (k - 1) / 2$ парних порівнянь.

Ідея тесту Тьюки — це використання для розрахунку t -величин не фактичних значень дисперсій даної пари групових середніх, а розраховану в процесі дисперсійного аналізу об'єднану внутрішньогрупову дисперсію ($MS_{wg} = MS_{error}$). Це дозволяє природно вирішити проблему сумісності множинних порівнянь дисперсійного аналізу з парними оцінками.

В процесі проведення апостеріорного парного аналізу ми отримуємо сукупність t -величин для всіх можливих пар вибірок. Їх розрахунок проводиться за розділом 5.1 глави 5.3.N.1. В позначеннях глави 13 рівняння (5.2 – 5.6) для довільної пари (ij) вибірок мають вигляд:

$$t_{ij} = \frac{|M_i - M_j|}{S_{d,ij}}. \quad (13.18)$$

$$v_{wg} = \sum_{i=1}^{i=k} (n_i - 1) = N - k. \quad (13.19)$$

$$p(> t_{ij}, v_{wg}) < \alpha. \quad (13.20)$$

При виконання нерівності (13.20), відмінність групових середніх є значущою на рівні α .

В тесті Тьюки замість співвідношень (5.2-5.3) використовують модифікований критерій Стьюдента:

$$s_{d,ij}^2 = \frac{MS_{wg}}{2} \times \left(\frac{1}{n_i} + \frac{1}{n_j} \right). \quad (13.21)$$

$$s_{d,ij} = \sqrt{s_{d,ij}^2}. \quad (13.22)$$

де MS_{wg} — об'єднана внутрішньогрупові дисперсія, яку розраховують в ході дисперсійного аналізу.

Завдяки тому, що в критерій (13.21) входить внутрішньогрупова дисперсія, забезпечується його сумісність з критеріями дисперсійного аналізу. Нульові гіпотези, що приймають або відкидають шляхом порівняння отриманих значень t -критерію (13.18) з критичним значенням для заданого рівня значущості α або розраховуючи відповідні ймовірності p_{ij} чи $P_{ij} \%$.

Якщо $p(t_{ij}, v) < \alpha = 0.05$ (5.0 %), то відмінність вважається значущою, а для $\alpha = 1.0 \%$ - високозначущою.

При оцінці отриманих t -величин використовують критерій найменшої значущої різниці (*LSD*), тобто дослідження зупиняють тоді, коли мінімальний критерій стає незначущим.

Відзначимо, що критерій Тьюки не дозволяє вирішити проблему росту ймовірності великих відмінностей з ростом числа вибірок. Для її рішення запропоновані різні критерії, з яких слід відзначити поправки Бонфероні і Холма-Бонфероні.

13.8.4. Поправки Бонфероні і Холма-Бонфероні

Ідея поправки Бонфероні така. Ми задаємо рівень значущості α для всієї сукупності парних t -критеріїв. Якщо число вибірок дорівнює k , то число парних комбінацій дорівнює:

$$m = k \times (k - 1) / 2. \quad (13.23)$$

Тому, для того, щоб рівень значущості α виконувався для всієї сукупності t -критеріїв, необхідно, щоб для кожного окремого t -критерію цей критерій дорівнював α / m . Відповідно, критерій (13.16) в підході Бонфероні має вигляд:

$$\text{Критерій Бонфероні: } p_{ij}(> t) \geq \alpha / m. \quad (13.24)$$

З співвідношення (13.24) видно, що в цьому випадку ймовірність того, що хоча б одне парне t -значення не буде відповідати вимогам (13.16) (тобто. вийде за рівень α), дорівнює $1 - [1 - (\alpha / m)]^m \approx \alpha$. Зокрема, для $k = 5$, $m = 10$, $\alpha = 0.05$ отримаємо: $1 - [1 - (0.05 / 10)]^{10} = 1 - 0.951 = 0.049 \approx 0.05$.

Неважко бачити, що з ростом числа вибірок k ймовірність ідентифікувати значущі відмінності між парами вибірок стає все менш і менш реальною. Тобто росте ймовірність хибнонегативних результатів — похибка другого роду (відмінність є статистично значущою, але не ідентифікується).

Цей недолік вирішує критерій Холма-Бонфероні. Основна його ідея: ймовірність (13.24) відноситься тільки до одного значення ймовірності. Після того, як ми її перевірили, кількість парних комбінацій зменшилось на одиницю. Відповідно, має змінитися і критерій (13.24). В загальному випадку, за послідовної схеми, описаній нижче, критерій Холма-Бонфероні має вигляд:

Критерій Холма-Бонфероні:

$$p_g(> t) \geq \alpha / (m - g + 1). \quad (13.25)$$

де g — номер ранжованої пари середніх (див. нижче).

Неважко бачити, що для першої ранжованої пари ($g = 1$) критерій Холма-Бонфероні (13.25) переходить в критерій Бонфероні (13.24), а для останньої ранжованої пари ($g = m$) — в звичайний t -критерій (13.16).

Таким чином, критерій Холма-Бонфероні (13.25) природно узагальнює t -критерій (13.16) і критерій Бонфероні (13.24), плавно змінюючись від великих до малих відмінностей.

Відзначимо, що критерій Холма-Бонфероні може бути застосовний як в підході Тьюки, так і в будь-якому іншому.

13.8.5. Процедура апостеріорного аналізу результатів методом парних порівнянь, що рекомендується

Процедура апостеріорного аналізу результатів методом парних порівнянь, що рекомендується, є поєднання тесту мінімальної значущої різниці (*LSD test*) і підходів Тьюки і Холма-Бонфероні.

Нехай ми маємо k середніх значень вибірок. Для них можливі $m = k \times (k - 1) / 2$ комбінацій пар середніх. Використовуючи підхід Тьюки (13.18-13.22), для кожної з них отримаємо t -значення, відповідні їм ймовірності p_{ij} , ранжуємо за збільшенням (від меншої до більшої) і присвоюємо їм номера $g = 1 \dots m$.

Для отриманої таким чином упорядкованої за збільшенням сукупності значень p_g проводимо процедуру *LSD* наступним чином:

Крок 1. Якщо $p_1 \geq \alpha / (m - 1 + 1)$, то приймають нульову гіпотезу рівності середніх (13.15) і подальший аналіз не проводять (оскільки інші пари тим більше будуть відповідати нульовій гіпотезі). Якщо $p_1 < \alpha / (m - 1 + 1)$, то нульову гіпотезу (13.15) відкидають для цієї пари і продовжують перевірку інших ранжованих ймовірностей на рівні значущості $\alpha / (m - 2 + 1)$.

Крок 2. Якщо $p_2 \geq \alpha / (m - 2 + 1)$, то приймають нульову гіпотезу і подальший аналіз не проводять. Якщо $p_2 < \alpha / (m - 2 + 1)$, то нульову гіпотезу відкидають і продовжують перевірку інших ранжованих ймовірностей на рівні значущості $\alpha / (m - 3 + 1)$.

І так далі...

Приклад застосування процедури, що рекомендується, представлений в Табл. 13.П-5.

13.9. МЕТОД КОНТРАСТІВ

Вище був розглянутий апостеріорний аналіз результатів методом парних порівнянь середніх за t -критерієм. Його завданням є ідентифікація тих пар середніх, які можна вважати (або не вважати) вибірками з однієї генеральної сукупності.

Одним з недоліків даного підходу є велика кількість (m) досліджуваних пар середніх. У загальному випадку, $m = k \times (k - 1) / 2$. Зокрема, якщо $k = 5$, то $m = 10$; якщо $k = 9$, то $m = 36$. Це створює певні труднощі при інтерпретації результатів. Багато з цих пар є лінійною комбінацією других пар і тому не несуть нової інформації у порівнянні з ними. Крім того, дослідника можуть цікавити інші завдання, наприклад:

- наскільки значущою є відмінність середніх від деякого середнього, який прийнято за базовий рівень (наприклад, результат аналізу за офіційною методикою);
- наскільки значущою є відмінність середніх від загального середнього значення всіх середніх;
- наскільки значущою є відмінність деяких двох сукупностей середніх однієї від іншої.

Можливі і інші завдання.

Такі можливості надає метод контрастів (*contrasts*). В ньому досліджуються не різниці пар середніх (як в методі парних порівнянь), а різниця (контраст) групових середніх і деякого базового рівня, який представляє інтерес для дослідника. Як такий може бути прийнятим одне з групових середніх або їх комбінація.

Відповідно, змінюється і постановка завдання. В методі контрастів шукають не мінімальну значущу різницю пар групових середніх (як в методі парних порівнянь), а оцінюють (звичайно за Стьюдентом) значущість відхилень групових середніх (або їх комбінації) від базового рівня. В цьому випадку в значній мірі втрачає актуальність проблема несумісності множинних парних порівнянь (див. 13.8.2).

Враховуючи різноманітність контрастів, їх застосування потребує чіткої постановки завдання дослідження. Без такої постановки інтерпретація отриманих результатів неможлива.

13.9.1. Визначення

Контраст — це зважена сума середніх значень кожного рівня якоїсь категоріальної змінної, в якій сума вагових коефіцієнтів звичайно (але не завжди) дорівнює нулю. В позначеннях глави 13, j -ий контраст (L_j) визначається таким чином:

$$L_j = \sum_{i=1}^{i=k} C_{ji} \times SM_i. \quad (13.26)$$

де C_{ji} — ваговий коефіцієнт j -ого контрасту для i -ого рівня досліджуваного фактора,
 SM_i — групове середнє для i -ого рівня досліджуваного фактора.

Співвідношення (13.26) можна записати також в матричному вигляді:

$$L = C \times SM. \quad (13.27)$$

де L — вектор контрастів,
 C — матриця контрастів,
 SM — вектор групових середніх.

В залежності від того, як задаються вагові коефіцієнти C_{ji} , контрасти дозволяють перевіряють різні гіпотези відносно групових середніх або їх комбінацій. Як вже відзначалося вище, звичайно (але не завжди) вагові коефіцієнти j -ого контрасту мають в сумі давати 0, тобто:

$$\sum_{i=1}^{i=k} C_{ji} = 0. \quad (13.28)$$

Для цього мають виконуватися наступні обов'язкові умови:

- коефіцієнти рівнів (або комбінацій рівнів), що порівнюються, мають мати різні знаки (+ або -),
- коефіцієнти рівнів, які не цікавлять, дорівнюють нулю.

Наприклад, якщо досліджуваний фактор має три рівня A , B і D , і ми хочемо «виявити контраст» (тобто різницю) між середніми значеннями рівнів A і D (тобто порівняти ці значення статистично), то коефіцієнтам контрасту $C_{A/D,i}$ можна надати такі значення:

$$\begin{matrix} C_{A/D,i} \\ A & -1 \\ B & 0 \\ D & 1 \end{matrix}$$

Аналогічно, при порівнянні середніх значень рівнів A і B коефіцієнти відповідного контрасту $C_{A/B,i}$ будуть мати такі значення:

$$\begin{matrix} C_{A/B,i} \\ A & -1 \\ B & 1 \\ D & 0 \end{matrix}$$

Вибір значень вагових коефіцієнтів в контрастах визначається поставленим завданням і є достатньо довільним — головне, щоб в сумі вони давали 0, тобто виконувалося співвідношення (13.28). Там, де це можливо, доцільно використовувати одиниці.

Відзначимо також, що кількість контрастів дорівнює $m = k - 1$, що значно менше, ніж комбінацій в методі парних порівнянь, який розглянутий вище ($m = k \times (k - 1) / 2$). В найпростішому випадку трьох рівнів ($k = 3$) ми можемо мати тільки два лінійно незалежних контрастів. Неважко бачити, що вектор контрасту $\langle C_{A/B} \rangle$ дорівнює різниці векторів $\langle C_{A/D} \rangle$ і $\langle C_{A/B} \rangle$. Тому при парному порівнянні середніх значень фактора с трьома рівнями ми отримуємо таку матрицю контрастів:

$$\begin{matrix} C_{A/D,i} & C_{A/B,i} \\ A & -1 & -1 \\ B & 0 & 1 \\ D & 1 & 0 \end{matrix}$$

За необхідності, ми можемо порівнювати також і комбінації з декількох рівнів фактора з другими його рівнями (або комбінаціями рівнів), що відкриває широкі можливості перевірки різних гіпотез по відношенню до даних, що аналізуються. Наприклад, якщо ми хочемо порівняти середнє значення рівня A з загальним середнім значенням рівнів B і D , то коефіцієнти такого контрасту можна представити таким чином:

$$\begin{matrix} C_{A/(B+D),i} \\ A & 2 \\ B & -1 \\ D & -1 \end{matrix}$$

Можна виділити чотири основних типів контрастів:

- контрасти комбінацій умов (*treatment contrasts*),
- контрасти сум (*sum contrasts*),
- контрасти користувача,
- контрасти Хелмерта (*Helmert contrasts*).

Метод контрастів передбачає використання розрахованої в ході дисперсійного аналізу об'єднаної вну-

трішньогрупової дисперсії MS_{wg} (див. розділ 13.8.3). Крім того, як правило, передбачається однаковий обсяг вибірки (n_i) і, відповідно, однакове число ступенів свободи (v_i) для всіх групових середніх. Виключенням є контрасти комбінацій умов, а також (в окремих випадках) контрасти користувача, які можуть застосовуватися і за різного обсягу вибірок групових середніх.

13.9.2. Контрасти комбінацій умов

Застосування цього типу контрастів передбачає, що один з рівнів досліджуваного фактору приймають за базовий (це можуть бути, наприклад, результати аналізу за якоюсь стандартною методикою), а всі інші рівні порівнюють з цим базовим рівнем.

Однією з переваг методу контрастів перед методом парних порівнянь є те, що перший використовує не самі середні значення, а їх відмінність від базового рівня. Це суттєво звужує діапазон варіювання результатів.

Застосування контрастів, в загальному випадку, передбачає однакове число ступенів свободи для всіх групових середніх. Однак контрасти комбінацій умов застосовні також і за різного числі ступенів свободи для досліджуваних групових середніх (як в Табл. 13.П-6). Цим пояснюється різниця в величинах $s_{d,ij}$ в Табл. 13.П-6 для різних контрастів (вони розраховуються за співвідношеннями (13.21-13.22)).

В Табл. 1.3.1 представлені принципи, які покладені в основу контрастів комбінацій умов.

Відзначимо, що перевірка нульової гіпотези за t -критерієм для параметра моделі β_0 (що сам базо-

вий рівень значуще відрізняється від нуля) у випадку кількісного визначення субстанцій (як в Табл. 13.П-1) є непотрібним.

Матриця контрастів комбінацій умов, яка побудована у відповідності з Табл. 13.1, представлена в Табл. 13.2.

Це — типовий приклад матриці, яка містить в собі вагові коефіцієнти «контрастів комбінацій умов». Як вже відзначалося вище, суми вагових коефіцієнтів за кожним стовбцем матриці контрастів не завжди дорівнюють нулю. Контрасти комбінацій умов - як раз той самий випадок (сума дорівнює 1), оскільки всі порівняння виконуються відносно базового рівня (групове середнє № 1 в розглянутому прикладі).

Контрасти комбінацій умов отримують за рівнянням (13.26) з використанням вектора групових середніх SM_i і матриці Табл. 13.2. Оцінка значущості відмінності отриманих стандартів від нуля проводиться за t -критерієм.

Для розрахунку $s_{d,ij}$, t -значень і $P(> |t|)$ використовують співвідношення (13.18-13.22), об'єднану внутрішньогрупову дисперсію MS_{wg} і об'єднане внутрішньогрупове число ступенів свободи v_{wg} .

Приклад розрахунку і оцінки контрастів комбінацій умов представлений в розділі 13.10.2.3.1 (Табл. 13.П-6).

13.9.3. Контрасти сум

При параметризації моделі дисперсійного аналізу за допомогою *контрастів сум* базовий рівень, з яким порівнюються інші рівні, представляє собою загальне середнє значення групових середніх (Y). Всі групові середні (Y_i) порівнюються з величиною Y .

Таблиця 13.1

Принципи контрастів комбінацій умов для базового рівня № 1

Параметр моделі	Оцінка	Нульова гіпотеза, що перевіряється
β_0	середнє значення для базового рівня: Y_1	$H_0: \mu_1 = 0$
β_1	середнє значення другого рівня за відрахуванням середнього значення базового рівня: $Y_2 - Y_1$	$H_0: \mu_2 - \mu_1 = 0$
β_2	середнє значення третього рівня за відрахуванням середнього значення базового рівня: $Y_3 - Y_1$	$H_0: \mu_3 - \mu_1 = 0$
β_3	середнє значення четвертого рівня за відрахуванням середнього значення базового рівня: $Y_4 - Y_1$	$H_0: \mu_4 - \mu_1 = 0$
β_4	середнє значення п'ятого за відрахуванням середнього значення базового рівня: $Y_5 - Y_1$	$H_0: \mu_5 - \mu_1 = 0$

Таблиця 13.2

Матриця контрастів комбінацій умов для базового рівня - № 1

№ рівня ↓	Контрасти				
	C_{21}	C_{31}	C_{41}	C_{51}	...
1 = базовий рівень	0	0	0	0	...
2	1	0	0	0	...
3	0	1	0	0	...
4	0	0	1	0	...
5	0	0	0	1	...
...

даному випадку ми з'ясуємо, наскільки значуще кожне з групових середніх відрізняється від загального середнього.

Відзначимо, що на відміну від контрастів комбінацій умов, обов'язковою вимогою застосування контрастів сум є однаковий обсяг вибірки для всіх середніх (i , відповідно, їх рівноточність), тобто:

$$n_i = n = \text{const}, i = 1 \dots k. \quad (13.29)$$

В Табл. 13.3 представлені принципи, які покладені в основу *контрастів сум*.

Дотримуючись Табл. 13.3, отримуємо матрицю контрастів сум, представлену в Табл. 13.4.

Контрасти комбінацій сум отримують за рівнянням (13.26) з використанням вектора групових середніх SM_i і матриці Табл. 13.4. Оцінка значущості відмінності отриманих контрастів від нуля проводиться за t -критерієм.

Для розрахунку t -значень і $P(>|t|)$ використовують співвідношення (13.18–13.22), об'єднану внутрішньогрупову дисперсію MS_{wg} і об'єднане внутрішньогрупове число ступенів свободи v_{wg} .

Певні труднощі в контрасті сум виникають з розрахунком величин $s_{d,ij}$, оскільки в цьому випадку до них не застосовні рівняння (13.21–13.22). Це пов'язано з тим, що групові середні Y_i і загальне середнє групових середніх $Y = \sum Y_j/k$ не є лінійно незалежними. Різницю $Y_i - Y$ можна представити у вигляді лінійно незалежних складових:

$$Y_i - Y = Y_i - \sum_{j=1}^{j=k} \frac{Y_j}{k} = \frac{k-1}{k} \times Y_i - \sum_{j \neq i} \frac{Y_j}{k} \quad (13.30)$$

Відповідно до (13.29), все групи мають однаковий обсяг n . Всі групи характеризуються об'єднаною внутрішньогруповою дисперсією одиничного MS_{wg} . Відповідно, об'єднана внутрішньогрупова дисперсія середнього дорівнює MS_{wg} / n . Тоді, проводячи від-

Таблиця 13.3

Принципи контрастів сум

Параметр моделі	Оцінка	Нульова гіпотеза, що перевіряється
β_0	середнє значення всіх групових середніх («загальне середнє»): $Y = \sum Y_i / k$	$H_0: Y = 0$
β_1	середнє значення першого рівня за відрахуванням загального середнього: $Y_1 - Y$	$H_0: Y_1 - Y = 0$
β_2	середнє значення другого рівня за відрахуванням загального середнього: $Y_2 - Y$	$H_0: Y_2 - Y = 0$
...
...

Таблиця 13.4

Матриця контрастів сум для випадку: $k = 5$, базовий рівень – середнє групових середніх

№ групи	Контрасти			
	C_1	C_2	C_3	C_4
1	1	0	0	0
2	0	1	0	0
3	0	0	1	0
4	0	0	0	1
5	-1	-1	-1	-1

Таблиця 13.5

Принципи контрастів Хелмерта

Параметр моделі	Оцінка	Нульова гіпотеза, що перевіряється
β_0	середнє значення з середніх по кожній групі («загальне середнє»): $Y = \sum Y_i / k$	$H_0: Y = 0$
β_1	$Y_2 - (Y_1 + Y)$	$H_0: Y_2 - (Y_1 + Y) = 0$
β_2	$Y_3 - (Y_1 + Y_2 + Y)$	$H_0: Y_3 - (Y_1 + Y_2 + Y) = 0$
...
...

Таблиця 13.6

Матриця контрастів Хелмерта для випадку: $k = 5$, базовий рівень – середнє середніх

	C_1	C_2	C_3	C_4
1	-1	-1	-1	-1
2	1	-1	-1	-1
3	0	2	-1	-1
4	0	0	3	-1
5	0	0	0	4

повідне диференціювання згідно співвідношення (9.3), отримуємо дисперсію середнього для величини $Y_1 - Y$:

$$\begin{aligned} MS(Y_1 - Y) &= s_{d,i}^2 = \\ &= \frac{(k-1)^2}{k^2} \times \frac{MS_{wg}}{n} + \frac{k-1}{k^2} \times \frac{MS_{wg}}{n} = \\ &= \frac{k-1}{n \times k} \times MS_{wg} \end{aligned} \quad (13.31)$$

Приклад розрахунку і оцінки контрастів сум наведений в розділі 13.10.2.3.2 (Табл. 13.П-7).

13.9.4. Контрасти користувача

Крім розглянутих вище стандартних типів контрастів, користувач може задавати свої власні контрасти, які призначаються для перевірки конкретних гіпотез щодо даних, що аналізуються.

Наприклад, в фармацевтичному аналізі часто виникає питання значущості різниці між двома сукупностями групових середніх. Цей тип контрастів є, фактично, об'єднанням *контрастів комбінацій умов і контрастів сум*. Розв'язання такого завдання зводиться до завдань об'єднання вибірок (див. розділ 1.3.1) і порівняння середніх результатів двох вибірок (див. розділ 5).

Приклад розрахунку такого типу контрастів користувача наведений в розділі 13.10.2.3.3 (Табл. 13.П-8).

13.9.5. Контрасти Хелмерта

При параметризації моделі дисперсійного аналізу за допомогою *контрастів Хелмерта* базовий рівень представляє собою загальне середнє значення з середніх по кожній групі («загальне середнє»). Порівняння групових середніх з цим базовим рівнем виконується за схемою, яка узагальнена в Табл. 13.5.

Як видно, суми за всіма стовбцями матриці контрастів Хелмерта дорівнюють нулю. Крім того, сума добутків елементів будь-яких двох стовбців матриці також дорівнює нулю. Контрасти з такими властивостями звуть *ортогональними*.

Контрасти Хелмерта вимагають чіткої постановки завдання. Тому, незважаючи на свої привабливі математичні властивості (ортогональність), вони рідко застосовуються на практиці, оскільки отримані параметри моделі складно інтерпретувати у застосуванні до конкретних проблема.

У цьому зв'язку, слід відзначити, що математика взагалі і статистика зокрема є потужним інструментом обробки і інтерпретації експериментальних даних, але вона вимагає чіткої постановки завдання і критеріїв прийнятності. У протилежному випадку отриманий в результаті статистичної обробки масив інформації є не тільки некорисним, але й створює непотрібний інформаційний шум.

13.10. ПРИКЛАДИ⁽¹⁾

13.10.1. Порівняння середніх результатів двох вибірок, отриманих при кількісному визначенні субстанції кислоти аскорбінової

При кількісному визначенні субстанції кислоти аскорбінової отримали результати, які представлені в Табл. 13.П-1. Для вибірок (груп) № 2 і 3 середні значення (Y), обсяги вибірок (n_i) і вибіркові стандартні відхилення (SD_i) мають вигляд:

$$\text{№ 2: } Y_2 = 99.56 \%, n_2 = 5, SD_2 = 0.188 \%,$$

$$\text{№ 3: } Y_3 = 99.19 \%, n_3 = 3, SD_3 = 0.026 \%.$$

Наскільки є значущим різниця цих вибіркових середніх?

Для відповіді на це питання використаємо t -критерій, який описаний в розділі 5. *Порівняння середніх результатів двох вибірок*. Із Табл. 13.П-4 видно, що різниця між SD_2 і SD_3 є статистично значущою за F -критерієм на рівні $1.95 \% < 5.0 \%$. Тому слід застосувати загальний підхід, який описаний в розділі 5.2. В позначеннях розділу 13 співвідношення (5.7-5.9) мають вигляд:

$$\begin{aligned} s_d^2 &= \frac{SD_2^2}{n_2} + \frac{SD_3^2}{n_3} = \frac{0.188^2}{5} + \frac{0.026^2}{3} = 0.007294. \\ s_d &= \sqrt{0.007294} = 0.085. \end{aligned}$$

$$\begin{aligned} v' &= (n_2 + n_3 - 2) \times \left(0.5 + \frac{SD_2^2 \times SD_3^2}{SD_2^4 + SD_3^4} \right) = \\ &= (5 + 3 - 2) \times \left(0.5 + \frac{0.188^2 \times 0.026^2}{0.188^4 + 0.026^4} \right) = \\ &= 6 \times 0.519 = 3.11. \end{aligned}$$

$$t = \frac{|Y_2 - Y_3|}{s_d} = \frac{99.56 - 99.19}{0.085} = 4.33.$$

$$P(t = 4.33, v' = 3.11) = 1.13 \% < 5.0 \% = \alpha.$$

Таким чином, ймовірність того, що групові середні $Y_2 = 99.56 \%$ і $Y_3 = 99.19 \%$ належить одній генеральній сукупності, складає менше 5.0% . Тобто відмінність між ними є значущим на рівні $\alpha = 5.0 \%$,

13.10.2. Кількісне визначення кислоти аскорбінової методом титрування: інтерпретація результатів

В Табл. 13.П-1 наведені результати (Y_{ij}) йодометричного кількісного визначення кислоти аскорбінової в субстанції. Аналіз проводили в $k = 5$ серіях дослідів (груп) з використанням різного аналітичного посуду і реактивів. Кількість паралельних визначень (n_i) в групах варіювалася від 3 до 7.

Постановка завдання:

(1) Розрахунки проводились з використанням комп'ютерних програм. Для проміжних етапів представлені округлені величини, тому можуть бути деякі розбіжності.

1. Чи є статистично значущими відмінності між середніми результатами аналізу різних груп?
2. Якщо так, то які пари середніх можна вважати вибірками з однієї генеральної сукупності, а які ні?
3. Середнє групи № 1 в Табл. 13.П-1 отримано в умовах, які визнані за стандартні. Наскільки значущими є відмінності інших групових середніх від цього середнього?
4. Наскільки є значущими відмінності групових середніх від загального середнього групових середніх?
5. Чи є значущими відмінності між сукупностями групових середніх?

Для вирішення *Завдання 1* використаємо однофакторний дисперсійний аналіз за загальною схемою, описаної в розділі 13.7.

Слід особливо підкреслити, що постановка *Завдань 2-5* має сенс тільки в тому випадку, якщо дисперсійний аналіз (*Задача 1*) покаже наявність значущих відмінностей між групами.

Для вирішення *Завдання 2* використаємо процедуру апостеріорного аналізу результатів методом парних порівнянь, описану в розділі 13.8.5.

Для вирішення *Завдань 3-5* використаємо методи контрастів:

Таблиця 13.П-1

Результати йодометричного кількісного визначення кислоти аскорбінової в субстанції в процентах ($k = 5$)

Вибірка (група) (i) → (j) ↓	1	2	3	4	5
n_i	5	5	3	7	5
$N = \sum n_i$	25				
	Y_{1j}	Y_{2j}	Y_{3j}	Y_{4j}	Y_{5j}
1	99.25	99.45	99.17	99.16	99.36
2	99.56	99.43	99.18	99.14	99.25
3	99.59	99.38	99.22	99.06	99.28
4	99.88	99.75		99.34	99.24
5	99.65	99.77		99.11	99.33
6				99.19	
7				99.29	
Внутрішньогрупові середні Y_i	99.59	99.56	99.19	99.18	99.29
Загальне середнє	$Y = 99.36$				
Внутрішньогрупові стандартні відхилення SD_i	0.226	0.188	0.026	0.099	0.052
Загальне стандартне відхилення	$SD_{tot} = 0.223$				
Внутрішньогрупові суми квадратів $SS_{wgi} = (n_i - 1) \times SD_i^2$	0.2041	0.1415	0.0014	0.0590	0.0107
Загальна внутрішньогрупова сума квадратів	$SS_{wg} = \sum SS_{wgi} = 0.417$				
Середня внутрішньогрупова сума квадратів	$MS_{wg} = SS_{wg} / (N - k) = 0.4167 / (25 - 5) = 0.0208$				
Загальна сума квадратів	$SS_{tot} = (N - 1) \times (SD_{tot})^2 = 24 \times 0.223^2 = 1.190$				
Середня загальна сума квадратів	$MS_{tot} = SS_{tot} / (N - 1) = 1.190 / (25 - 1) = 0.0496$				
Міжгрупова сума квадратів	$SS_{effect} = SS_{bg} = SS_{tot} - SS_{wg} = 1.190 - 0.417 = 0.773$				
Середня міжгрупова сума квадратів	$MS_{bg} = SS_{bg} / (k - 1) = 0.773 / (5 - 1) = 0.193$				
$F =$	$MS_{bg} / MS_{wg} = 0.193 / 0.0208 = 9.28$				
$P(F, k - 1, N - k) = P(9.28, 4, 20)$	$0.021 < 1.0 \%$				
Висновок:	Відмінність групових середніх є високозначущою				
Коефіцієнт детермінації η^2	$100 \times SS_{effect} / SS_{tot} = 100 \times 0.773 / 1.190 = 65.0 \%$				

Таблиця 13.П-2

Таблиця дисперсійного аналізу для результатів йодометричного кількісного визначення кислоти аскорбінової

	Головний ефект				
	ν	SS	MS	F	$P(F) \%$
Міжгрупова (bg) варіація ($effect$)	$k - 1 = 4$	0.773	0.193	9.28	$0.021 < 1.0 \%$
Внутрішньогрупова (wg) варіація ($error$)	$N - k = 20$	0.417	0.0208		
В цілому (tot)	$N - 1 = 24$	1.190	0.0496		
Коефіцієнт детермінації η^2					0.650 або 65.0 %

Завдання 3 — контраст комбінацій умов, описаний в розділі 13.9.2.

Завдання 4 — контраст сум, описаний в розділі 13.9.3.

Завдання 5 — контраст користувача, описаний в розділі 13.9.4.

13.10.2.1. Дисперсійний аналіз результатів

Постановка Завдання 1

Чи є статистично значущими відмінності між середніми результатами аналізу різних груп?

Для розв'язання цього завдання використаємо загальну схему дисперсійного аналізу, описану в розділі 13.7.

Перевірка гіпотези однорідності внутрішньогрупових дисперсій (13.1).

Ця процедура є обов'язковою вимогою для застосування дисперсійного аналізу (див. розділ 13.4). Враховуючи різні обсяги групових вибірок, доцільно використати критерій Левене (див. розділ 1.3.4).

Результати порівняння внутрішньогрупових дисперсій за критерієм Левене представлені в Табл. 13.П-3. Ймовірність того, що вони є вибірками з однієї генеральної сукупності, складає $P(>F = 2.69, 4, 20) = 6.06 > 5.0 \%$, тобто відмінність дисперсій є незначущою на рівні 5.0 %. Таким чином, гіпотеза однорідності дисперсій (13.1) підтверджується.

Одночасно, в Табл. 13.П-4 показано, що хоча однорідність групових дисперсій і виконується за критерієм Левене, але з 10 пар групових дисперсій для 4 пар відмінності є статистично значущими, а для 2 пар — високозначущими за критерієм Фішера. Це ілюструє несумісність парних і групових критеріїв і необхідність використання для дисперсійного аналізу об'єднаної внутрішньогрупової дисперсії MS_{wg} .

Вихідні результати і розрахунок дисперсійних характеристик за допомогою співвідношень (13.3-13.6) представлені в Табл. 13.П-1.

Результати розрахунку дисперсійних характеристик Табл. 13.П-1 зручно підсумувати в Табл. 13.П-2, яку прийнято звати таблицею дисперсійного аналізу.

Як видно з Табл. 13.П-2, відмінність групових середніх є високозначущою ($P(F) = 0.021 < 1.0 \%$).

Високий коефіцієнт детермінації (65.0 %) свідчить про значний вплив систематичної похибки на результати йодометричного аналізу кислоти аскорбінової. Це узгоджується з загальними рекомендаціями при валідації титрометричних методик (див. розділ 9.7.1 глави 5.3.N.2 *Валідація аналітичних методик і випробувань*). Причинами є систематичні похибки, пов'язані з пробопідготовкою, визначенням титру і іншими операціями, які не компенсуються в прямих методах, на відміну від метода стандарту.

Оскільки відмінність групових середніх є високозначущою, то ми можемо перейти до *Завдань 2-5*: — апостеріорний аналіз результатів методом парних порівнянь (див. розділ 13.8.5), — методи контрастів (див. розділ 13.9.2-13.9.4).

13.10.2.2. Парні порівняння

Постановка Завдання 2

Які пари групових середніх можна вважати вибірками з однієї генеральної сукупності, а які ні?

В Табл. 13.П-5 представлено застосування процедури апостеріорного аналізу методом парних порівнянь, яка рекомендована в п. 13.7.5. В розрахунках використані результати Табл. 13.П-1.

Приклад розрахунку для пари порівнянь групових середніх 2/3:

$$Y_2 = 99.556, n_2 = 5.$$

$$Y_3 = 99.190, n_3 = 5.$$

$$|Y_2 - Y_3| = 99.556 - 99.190 = 0.366.$$

$$MS_{wg} = 0.0208, v = 20,$$

$$m = k \times (k - 1) = 5 \times 4 / 2 = 10.$$

$$s_{d,23}^2 = 0.0208 \times \left(\frac{1}{5} + \frac{1}{5} \right) = 0.0102.$$

$$s_{d,23} = \sqrt{0.0102} = 0.101.$$

$$t_{23} = 0.366 / 0.101 = 3.472.$$

$P(t_{23} = 3.472, v = 20) = 0.120 \%$. Ранжування: $g = 4$ (см. Табл. 13-П.5).

Критерій Холма-Бонфероні для $\alpha = 0.05, m = 10, g = 4$:

$$P_g = \alpha / (m - g + 1) = 0.05 / (10 - 4 + 1) = 0.05 / 7 = 0.00714 \text{ або } 0.714 \%.$$

$$P(t_{23} = 3.472, v = 20) = 0.120 < P_g = 0.714 \%.$$

Співвідношення (13.25) не виконується, тому рухаємося далі за процедурою, яка описана в розділі 13.8.5.

Висновок:

З Табл. 13.П-5 видно, що критичною (яка відповідає принципу *LSD*) є пара 4/5. Відповідно, пари 4/5, 3/5, 1/2 і 3/4 належать до генеральної сукупності на рівні значущості 5.0%. Інші 6 пар не належать до генеральної сукупності на рівні значущості 5.0%.

Причинами, як вже обговорювалося раніше, є систематичні похибки, пов'язані з відмінністю бюреток, титрантів і т.д. в різних серіях дослідів. Це є характерним для прямих методів, до яких належить титрометрія.

Перевірка однорідності дисперсій за критерієм Левене результатів йодометричного кількісного визначення кислоти аскорбінової ($k = 5$)

Назва	Група 1			Група 2			Група 3			Група 4			Група 5			Сума
$n_i =$		5			5			7			3			5		25 = N
$j \downarrow$	Y_{ij}	$Z_{ij}^{(1)}$	$(Z_{ij}^{(1)} - Z_i)^2$	Y_{2j}	Z_{2j}	$(Z_{2j} - Z_2)^2$	Y_{3j}	Z_{3j}	$(Z_{3j} - Z_3)^2$	Y_{4j}	Z_{4j}	$(Z_{4j} - Z_4)^2$	Y_{5j}	Z_{5j}	$(Z_{5j} - Z_5)^2$	
1	99.25	0.34	0.0366	99.45	0.11	0.0033	99.17	0.02	0.0000	99.16	0.02	0.0027	99.36	0.07	0.0007	
2	99.56	0.03	0.0141	99.43	0.13	0.0014	99.18	0.01	0.0001	99.14	0.04	0.0010	99.25	0.04	0.0000	
3	99.59	0.00	0.0198	99.38	0.18	0.0002	99.22	0.03	0.0001	99.06	0.12	0.0023	99.28	0.01	0.0009	
4	99.88	0.29	0.0223	99.75	0.19	0.0009				99.34	0.16	0.0063	99.24	0.05	0.0001	
5	99.65	0.06	0.0065	99.77	0.21	0.0026				99.11	0.07	0.0000	99.33	0.04	0.0000	
6										99.19	0.01	0.0050				
7										99.29	0.11	0.0009				
Групові середні Y_i	99.59			99.56			99.19			99.18			99.29			
Групові суми Σ_i		0.72			0.82			0.06			0.53			0.21		2.346
Групові середні $Z_i = \Sigma_i / n_i$		0.14			0.16			0.02			0.08			0.04		
Загальне середнє $Z = \Sigma_i / N = 2.346 / 25 = 0.0939$																
Групові суми Σ_i			0.0993			0.0083			0.0002			0.0182			0.0017	0.1277 ²⁾
$n_i \times (Z_i - Z)$		0.0130			0.0240			0.0164			0.0021			0.0132		0.0688 ³⁾
Суми квадратів SS_{wg}		0.2041			0.1415			0.0014			0.0590			0.0107		0.4167 ⁴⁾
$F = W = [(N - k) / (k - 1)] \times \text{чисельник} / \text{знаменник} =$ $= [(25 - 5) / (5 - 1)] \times 0.0688 / 0.1277 = 2.69. P(F = 2.69, 4, 20) = 6.06 > 5.0\%.$																
Висновок: відмінність дисперсій є незначущою на рівні 5.0 %																

1) — $Z_{ij} = |Y_{ij} - Y_i|$

2) — знаменник

3) — чисельник

4) — загальна внутрішньогрупова сума квадратів SS_{wg}

Таблиця 13.П-4

Значущість відмінностей групових дисперсій за Фішером для різних пар групових середніх

№ групи	Обсяг вибірки n_i	Внутрішньогрупові SD_i	Пари середніх (ij)	$F_{ij} = (SD_i / SD_j)^2$ або $F_{ji} = (SD_j / SD_i)^2$	$P(>F_{ij}, n_i - 1, n_j - 1) \%$ або $P(>F_{ji}, n_j - 1, n_i - 1) \%$
1	5	0.226	1/2	1.44	36.57
2	5	0.188	1/3	72.9	1.36*
3	3	0.026	1/4	5.19	3.75*
4	7	0.099	1/5	19.1	0.72**
5	5	0.052	2/3	50.5	1.95*
			2/4	3.60	7.93
			2/5	13.3	1.41*
			3/4	14.0	0.55**
			3/5	3.81	11.83

* $P < 5.0\%$: відмінність дисперсій є значущою,

** $P < 1.0\%$: відмінність дисперсій є високозначущою

13.10.2.3. Контрасти

13.10.2.3.1. Контраст комбінацій умов

Постановка Завдання 3

Середнє групи № 1 в Табл. 13.П-1 отримано в умовах, які прийняті за стандартні. Наскільки значуще відрізняються інші групові середні від цього середнього?

Використаємо результати, представлені в Табл. 13.П-1. Як базовий рівень виберемо середнє № 1 (99.59 %). Для розрахунків величин $s_{d,ij}$, t -значень і $P(>|t|)$ використаємо співвідношення (13.18-13.22), об'єднану внутрішньогрупову дисперсію $MS_{wg} = 0.0208$, а також фактичні обсяги вибірок n_i (див. Табл. 13.П-4). Розрахунки проводимо за схемою, яка описана вище в розділі 13.10.2.2. Результати розрахунків контрастів комбінацій умов представлені в Табл. 13.П-6.

Неважко бачити, що контрасти комбінацій умов, фактично, являють собою спрощену версію метода парних порівнянь (див. розділи 13.7-13.8). Замість 10 пар середніх, у нас є тільки 4 контраста, що суттєво спрощує інтерпретацію результатів.

З Табл. 13.П-6 слідує, що групові середні № 3-5 високозначуще ($P < 1.0\%$) відрізняються від групового середнього № 1. Різниця ж групового середнього № 2 ($P = 37.3\%$) від групового середнього № 1 є статистично незначущим на рівні $> 5.0\%$.

13.10.2.3.2. Контрасти сум

Постановка Завдання 4

Наскільки значущими є відмінності групових середніх від загального середнього групових середніх?

Як базовий рівень беремо загальне середнє групових середніх 99.36 % (див. Табл. 13.П-1). Для розрахунків t -значень використаємо співвідношення (13.18-13.20) і об'єднану внутрішньогрупову дисперсію $MS_{wg} = 0.0208$.

Розрахунок величин $s_{d,ij}$ проводимо за співвідношенням (13.31), використовуючи результати Табл. 13.9-1: $MS_{wg} = 0.0208$, $k = 5$. Відповідно до вимог (13.29), приймемо для нашого випадку $n = N/k = 25/5 = 5$. Тоді співвідношення (13.31) дає:

$$s_{d,i}^2 = \frac{k-1}{n \times k} \times MS_{wg} = \frac{5-1}{5 \times 5} \times 0.0208 = 0.003334.$$

$$s_{d,i} = \sqrt{0.003334} = 0.0577.$$

Розрахунок t -значень проводимо за рівнянням (13.18-13.20), використовуючи отримане значення $s_{d,i}$. Результати розрахунків представлені в Табл. 13.П-7.

Як описано вище в розділі 13.8.1, в Табл. 13.П-7 мають сенс тільки чотири порівняння (контрасти) з загальним середнім значенням. П'ятий контраст (який наданий тільки для ілюстрації) є лінійною комбінацією 4 попередніх контрастів і не несе, тому, нової інформації. Зокрема, він (-0.070) дорівнює сумі контрастів 1-4 (0.070), взятої з оборотним знаком.

Із Табл. 13.П-7 видно, що всі вибрані групові середні (№ 1-4) на рівні $< 1.0\%$ (тобто високозначуще) відрізняються від загального середнього.

Таблиця 13.П-5
Апостеріорний аналіз результатів методом парних порівнянь ($k = 5$, $m = 10$, $v = 20$, $MS_{wg} = 0.0208$)

№	Середні Y_i	n_i	Пари порівнянь (ij)	$ Y_i - Y_j $	$s_{d,ij}$	t_{ij}	$P(t_{ij}, v = 20)$ %	Пари порівнянь (ij)	Ранжовані P_{ij}	Ранжування, g	Критерій Холма-Бонфероні
1	99.586	5	1/2	0.030	0.091	0.329	37.29	1/4	0.00608	1	0.500
2	99.556	5	1/3	0.396	0.105	3.757	0.0621	2/4	0.0139	2	0.556
3	99.190	3	1/4	0.402	0.085	4.753	0.00608	1/3	0.0621	3	0.625
4	99.184	7	1/5	0.294	0.091	3.221	0.215	2/3	0.120	4	0.714
5	99.292	5	2/3	0.366	0.105	3.472	0.120	1/5	0.215	5	0.833
			2/4	0.372	0.085	4.398	0.0139	2/5	0.451	6	1.000
			2/5	0.264	0.091	2.892	0.451	4/5	10.86	7	1.250
			3/4	0.006	0.100	0.057	47.74	3/5	17.24	8	1.667
			3/5	0.102	0.105	0.968	17.24	1/2	37.29	9	2.500
			4/5	0.108	0.085	1.274	10.86	3/4	47.74	10	5.000

Таблиця 13.П-6

Розрахунок контрастів комбінацій умов для результатів аналізу кислоти аскорбінової (базовий рівень № 1: $Y_1 = 99.59$, $MS_{wg} = 0.0208$)

Номер рівня (групи)	Контрасти у порівнянні з базовим рівнем № 1	$s_{d,i}$	t -значення	$P(> t)\%$
2	0.030	0.0913	0.329	37.3
3	0.396	0.1054	3.76	0.0621
4	0.402	0.0845	4.75	0.00608
5	0.294	0.0913	3.22	0.215

13.10.2.3.3. Контрасти користувача

Постановка Завдання 5

Аналіз даних Табл. 13.П-1 показує, що можна припустити значущу різницю між груповими середніми $Y_1 = 99.59\%$ ($n_1 = 5$) і $Y_2 = 99.56\%$ ($n_2 = 5$), з однієї сторони, і $Y_3 = 99.19\%$ ($n_3 = 3$), $Y_4 = 99.18\%$ ($n_4 = 7$) і $Y_5 = 99.29\%$ ($n_5 = 5$), з іншої сторони. Чи є статистично значущою різниця між цими двома сукупностями групових середніх?

Знайдемо середні значення і ступені свободи, в позначеннях розділу 13, для цих двох сукупностей, використовуючи формулу (1.16):

$$Y_{1,2} = (5 \times 99.59 + 5 \times 99.56) / (5+5) = 99.57\%,$$

$$v_{1,2} = (5 - 1) + (5 - 1) = 8,$$

$$Y_{3,4,5} = (3 \times 99.19 + 7 \times 99.18 + 5 \times 99.29) / (3 + 7 + 5) = 99.22\%,$$

$$v_{3,4,5} = (3 - 1) + (7 - 1) + (5 - 1) = 12.$$

Критерій Стюдента розраховуємо за співвідношенням (5.4), покладаючи $s_p = \sqrt{MS_{wg}} = \sqrt{0.0208} = 0.144$:

$$t = \frac{99.57 - 99.22}{0.144} \times \sqrt{\frac{(5+5) \times (3+7+5)}{(5+5) + (3+7+5)}} = 5.93.$$

$$P(5.93, 8, 12) = 0.322\% < 1.0\%.$$

Як видно, різниця між цими двома сукупностями групових середніх є високозначущою.

13.10.3. Відпрацювання технології отримання таблеток: взаємозв'язок між однорідністю маси і дозування

В процесі відпрацювання технології отримання таблеткової лікарської форми деякого нового препарату вивчали вплив невеликих варіацій технологічних параметрів на середню масу таблеток, кількісне визначення і фармакопейне випробування на однорідність вмісту (2.9.40).

Перед дослідженнями ставились наступні завдання:

1. Наскільки значущим є вплив невеликих варіацій в технології отримання таблеток на середню масу, кількісне визначення, однорідність маси (UM) і однорідність вмісту (UC)?

2. Чи існує для даної технології виробництва препарату залежність між кількісним визначенням і середньою масою?
3. Чи існує для даної технології виробництва препарату залежність між однорідністю маси і однорідністю вмісту?

Для відповіді на ці питання досліджували по 10 таблеток з 9 різних серій (груп), отриманих з деякими варіаціями технологічних параметрів. Для кожної таблетки отримували значення маси (UM_{ij}) і кількісного визначення (UC_{ij}) в відсотках до номінальних значень, групові середні (SM), а також відношення $100 \times UC_{ij} / UM_{ij} \%$. З отриманих величин розраховували середнє для кожної серії ($100 \times UC_{ij} / UM_{ij}$) і стандартне відхилення, а також загальне середнє для кожної серії і загальне стандартне відхилення.

Результати досліджень представлені в Табл. 13.П-8.

Відповідь на перше питання дозволяє дати однофакторний дисперсійний аналіз, приклад застосування якого описаний вище в розділі 13.10.2.1. Для відповіді на друге і третє питання зручно застосувати регресійний аналіз за допомогою лінійного методу найменших квадратів (ЛМНК), приклад застосування якого описаний в розділі 12.9.1.

13.10.3.1. Дисперсійний аналіз однорідності маси

Використаємо підхід, описаний в Табл. 13.П-1.

Кількість груп (рівнів, серій):
 $k = 9$.

Обсяги вибірок: $n_i = n = 10$.

Загальний обсяг всіх вибірок:
 $N = k \times n = 9 \times 10 = 90$.

Ступені свободи:
 $v_{bg} = k - 1 = 9 - 1 = 8$.
 $v_{wg} = N - k = 90 - 9 = 81$.
 $v_{tot} = N - 1 = 90 - 1 = 89$.

Перевірка однорідності групових дисперсій

Ця перевірка є обов'язковою перед проведенням дисперсійного аналізу (див. розділ 13.4). Оскільки в Табл. 13.П-8 представлені тільки стандартні відхилення (а не самі внутрішньогрупові значення функ-

Таблиця 13.П-7

Приклад розрахунку контрастів сум для результатів аналізу кислоти аскорбінової ($N = 25, k = 5, t = k - 1 = 4$)

Базовий рівень	$Y = (99.59 + 99.56 + 99.19 + 99.18 + 99.29)/5 = 99.36$	$MS_{wg} = 0.0208$		
Номер рівня	Контрасти порівняно з базовим рівнем	$s_{d,ij}$	t-значення	$P(> t) \%$
1	0.225	0.0577	3.89	0.046
2	0.195	0.0577	3.37	0.154
3	-0.171	0.0577	2.97	0.376
4	-0.177	0.0577	3.07	0.301
Сума 1-4	0.070			
5	-0.070	0.0577	1.21	12.1

ції відгуку) і обсяги вибірок однакові ($n_i = n = 10$), то зручно застосувати простий критерій Кокрена (1.19) (див. розділ 1.3.3). Це співвідношення в термінах глави 13 має вигляд:

$$G = \frac{SD_{max}^2}{\sum_{i=1}^{k-1} SD_i^2} \quad (13.П-1)$$

З Табл. 13.П-8 видно, що $SD_{max} = 2.14$. Використовуючи дані Табл. 13.П-8, із співвідношення (13.П-1) отримаємо:

$$G = 2.14^2 / (2.14^2 + 1.31^2 + 1.19^2 + 1.50^2 + 1.44^2 + 0.95^2 + 1.54^2 + 1.09^2 + 1.32^2) = 0.25.$$

Відповідно до Табл. 10.4, величина G не має перевищувати $G(95\%, 9, 9) = 0.27$. Як видно $0.25 < 0.27$, тому групові дисперсії можна вважати однорідними.

Внутрішньогрупова варіація:

Середня сума квадратів (дисперсія):

$$MS_{wg} = (SD_{wg})^2 = 1.42^2 = 2.03$$

Сума квадратів:

$$SS_{wg} = MS_{wg} \times v_{wg} = 2.03 \times 81 = 164.03$$

Загальна варіація:

Середня сума квадратів (дисперсія):

$$MS_{tot} = (SD_{tot})^2 = 1.63^2 = 2.66$$

Сума квадратів:

$$SS_{tot} = MS_{tot} \times v_{tot} = 2.66 \times 89 = 236.98$$

Міжгрупова варіація:

Сума квадратів:

$$SS_{bg} = SS_{tot} - SS_{wg} = 236.98 - 164.03 = 72.95$$

Середня сума квадратів (дисперсія):

$$MS_{bg} = SS_{bg} / v_{bg} = 72.95 / 8 = 9.12$$

Співвідношення Фішера:

$$F = MS_{bg} / MS_{wg} = 9.12 / 2.03 = 4.50$$

Ймовірність бути вибітками з однієї генеральної сукупності

$$P(>F = 4.50, v_{bg} = 8, v_{wg} = 81) = 0.015\%$$

Коефіцієнт детермінації:

$$\eta^2 = 100 \times SS_{bg} / SS_{tot} = 100 \times 72.95 / 236.98 = 30.8\%$$

Результати розрахунків підсумовані в таблиці дисперсійного аналізу (Табл. 13.П-9).

Як видно, міжгрупові відмінності високозначуще більше, ніж внутрішньогрупові. Відмінності ці настільки великі, що не має необхідності проводити апостеріорний аналіз з метою з'ясування причин цих відмінностей.

Таким чином, варіація технологічних параметрів призводить до високозначущої відмінності серій за показником середньої маси таблеток.

13.10.3.2. Дисперсійний аналіз однорідності вмісту

Кількість груп (рівнів, серій):

$$k = 9.$$

Обсяги вибірок:

$$n_i = n = 10.$$

Загальний обсяг всіх вибірок:

$$N = k \times n = 9 \times 10 = 90.$$

Ступені свободи:

$$v_{bg} = k - 1 = 9 - 1 = 8.$$

$$v_{wg} = N - k = 90 - 9 = 81.$$

$$v_{tot} = N - 1 = 90 - 1 = 89.$$

Перевірка однорідності групових дисперсій

Таблиця 13.П-8

Результати дослідження експериментальних серій таблеток на однорідність маси і вмісту

№ серії	Однорідність маси $UM\%$		Однорідність вмісту $UC\%$		$(100 \times UC_{ij} / UM_{ij})\%$	
	SM_i	SD_i	SM_i	SD_i	SM_i	SD_i
1	102.28	2.14	106.20	2.38	103.83	0.63
2	100.28	1.31	103.27	1.51	102.98	0.73
3	101.41	1.19	103.42	1.42	101.98	0.33
4	102.43	1.50	105.85	1.80	103.34	0.44
5	100.34	1.44	103.77	1.50	103.42	0.36
6	100.12	0.95	102.20	0.94	102.08	0.26
7	102.04	1.54	105.50	1.90	103.39	0.40
8	100.25	1.09	102.81	1.02	102.56	0.40
9	100.90	1.32	103.53	1.54	102.60	0.41
Загальні середні SM	101.12		104.06		102.91	
Загальне стандартне відхилення SD_{tot}	1.63		2.06		0.75	
Об'єднане внутрішньогрупове станд. відхилення SD_{wg}		1.42		1.61		0.46
Стандартне відхилення групових середніх SD_{SM}	0.94		1.43		0.64	
Стандартні відхилення групових дисперсій*		1.07		1.45		0.15

* $[(\sum SD_i^2) / (9 - 1)]^{0.5}$

З Табл. 13.П-8 видно, що $SD_{max} = 2.38$. Використовуючи дані Табл. 13.П-8 і слідуючи за процедурою, яка описана вище в розділі 13.10.3.1, отримаємо:

$$G = 2.38^2 / (2.38^2 + 1.51^2 + 1.42^2 + 1.80^2 + 1.50^2 + 0.94^2 + 1.90^2 + 1.02^2 + 1.54^2) = 0.24 < G(95\%, 9, 9) = 0.27.$$

Як видно групові дисперсії можна вважати однорідними.

Внутрішньогрупова варіація:

Середня сума квадратів (дисперсія):

$$MS_{wg} = (SD_{wg})^2 = 1.61^2 = 2.60$$

Сума квадратів:

$$SS_{wg} = MS_{wg} \times v_{wg} = 2.60 \times 81 = 210.56$$

Загальна варіація:

Середня сума квадратів (дисперсія):

$$MS_{tot} = (SD_{tot})^2 = 2.06^2 = 4.24$$

Сума квадратів:

$$SS_{tot} = MS_{tot} \times v_{tot} = 2.75 \times 89 = 377.25$$

Міжгрупова варіація:

Сума квадратів:

$$SS_{bg} = SS_{tot} - SS_{wg} = 377.25 - 210.56 = 166.69$$

Середня сума квадратів (дисперсія):

$$MS_{bg} = SS_{bg} / v_{bg} = 166.69 / 8 = 20.84$$

Співвідношення Фішера:

$$F = MS_{bg} / MS_{wg} = 20.84 / 2.60 = 8.02$$

Ймовірність бути вибірками з однієї генеральної сукупності:

$$P(>F = 8.02, v_{bg} = 8, v_{wg} = 81) = 0.0000067\%$$

Коефіцієнт детермінації:

$$\eta^2 = 100 \times SS_{bg} / SS_{tot} = 100 \times 166.69 / 377.25 = 44.2\%$$

Результати розрахунків підсумовані в таблиці дисперсійного аналізу (Табл. 13.П-10).

Як видно, міжгрупові відмінності високозначуще більше, ніж внутрішньогрупові. Як і для однорідності маси, ці відмінності настільки великі, що не має необхідності проводити апостеріорний аналіз з метою з'ясування причин цих відмінностей..

Таким чином, варіація технологічних параметрів призводить до високозначущій різниці серій за показником кількісного визначення.

13.10.3.3. Дисперсійний аналіз величини

$$100 \times UC_{ij} / UM_{ij} \%$$

Кількість груп (рівнів, серій):

$$k = 9.$$

Обсяги вибірок:

$$n_i = n = 10.$$

Загальний обсяг всіх вибірок:

$$N = k \times n = 9 \times 10 = 90.$$

Ступені свободи:

$$v_{bg} = k - 1 = 9 - 1 = 8.$$

$$v_{wg} = N - k = 90 - 9 = 81.$$

$$v_{tot} = N - 1 = 90 - 1 = 89.$$

Перевірка однорідності групових дисперсій

З Табл. 13.П-8 видно, що $SD_{max} = 0.73$. Використовуючи дані Табл. 13.П-8 і слідуючи за процедурою, яка описана вище в розділі 13.10.3.1, отримаємо:

$$G = 0.73^2 / ((0.63^2 + 0.73^2 + 0.33^2 + 0.44^2 + 0.36^2 + 0.26^2 + 0.40^2 + 0.40^2 + 0.41^2)) = 0.27 \leq G(95\%, 9, 9) = 0.27.$$

Як видно групові дисперсії можна вважати однорідними.

Внутрішньогрупова варіація:

Середня сума квадратів (дисперсія):

$$MS_{wg} = (SD_{wg})^2 = 0.46^2 = 0.21$$

Сума квадратів:

$$SS_{wg} = MS_{wg} \times v_{wg} = 0.21 \times 81 = 17.28$$

Загальна варіація:

Середня сума квадратів (дисперсія):

$$MS_{tot} = (SD_{tot})^2 = 0.75^2 = 0.57$$

Сума квадратів:

$$SS_{tot} = MS_{tot} \times v_{tot} = 0.57 \times 89 = 50.63$$

Міжгрупова варіація:

Сума квадратів:

$$SS_{bg} = SS_{tot} - SS_{wg} = 50.63 - 17.28 = 33.34$$

Середня сума квадратів (дисперсія):

$$MS_{bg} = SS_{bg} / v_{bg} = 33.34 / 8 = 4.17$$

Співвідношення Фішера:

$$F = MS_{bg} / MS_{wg} = 4.17 / 0.21 = 19.53$$

Ймовірність бути вибірками з однієї генеральної сукупності:

$$P(>F = 8.02, v_{bg} = 8, v_{wg} = 81) = 0.00\%$$

Таблиця 13.П-9

Таблиця дисперсійного аналізу для однорідності маси

	Головний ефект				
	<i>v</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P(>F) %</i>
Міжгрупова (<i>bg</i>) варіація (<i>effect</i>)	8	72.95	9.12	4.50	0.015 < 1.0 %
Внутрішньогрупова (<i>wg</i>) варіація (<i>error</i>)	81	164.03	2.03		
В цілому (<i>tot</i>)	89	236.98	2.66		
Коефіцієнт детермінації η^2					0.308 або 30.8 %

Коефіцієнт детермінації:

$$\eta^2 = 100 \times SS_{bg} / SS_{tot} = 100 \times 33.34 / 50.63 = 65.9 \%$$

Результати розрахунків підсумовані в таблиці дисперсійного аналізу (Табл. 13.П-11).

Як видно, міжгрупові відмінності високозначуще більше, ніж внутрішньогрупові. Як і в двох попередніх випадках, відмінності ці настільки великі, що не має необхідності проводити апостеріорний аналіз з метою з'ясування причин цих відмінностей.

Таким чином, використання відношення кількісного визначення в одній таблетці до її маси не зменшує вплив технологічних параметрів. Відмінності між серіями залишаються високозначущими.

13.10.3.4. Регресійний аналіз взаємозв'язку між однорідністю вмісту і однорідністю маси

13.10.3.4.1. Вихідні положення

Досліджувані групові середні варіюють в достатньо вузьких межах, тому можна чекати, що кореляція між ними, якщо вона є, має лінійний характер, тобто.

$$y = b \times x + a. \quad (13.П-2)$$

Після обробки експериментальних даних за допомогою ЛМНК за схемою, яка описана в розділі 12.9.1, ми отримуємо величини b і a , їх стандартні відхилення s_b і s_a , залишкове стандартне відхилення s_0 і загальний індекс (коефіцієнт) кореляції R_c . Ці метрологічні величини дозволяють оцінити статистичну значущість лінійної залежності (13.П-2).

В залежності від поставленого завдання, величини y і x можуть бути різними. В даному випадку нас цікавить:

- 1) Залежність групових середніх результатів кількісного визначення в одній таблетці в межах проведення фармакопейного випробування (2.9.40) на однорідність вмісту ($SM(UC)$) від середньої маси цих же таблеток $SM(UM)$ – див. Табл. 13.П-8.

Враховуючи невеликі відхилення від 100% досліджуваних величин, доцільно дослідити взаємозв'язок не між самими величинами, а між їх відхиленнями від 100%. На абсолютні метрологічні характеристики прямої це не впливає (оскільки це просто вибір системи координат), але суттєво спрощує розрахунки і робить їх більш наочними.

Таким чином, в цьому випадку в рівнянні (13.П-1) маємо:

$$y = SM(UC) - 100; \quad x = SM(UM) - 100.$$

- 2) Залежність дисперсії однорідності вмісту в межах проведення фармакопейного випробування (2.9.40) на однорідність вмісту ($SD^2(UC)$) від дисперсії маси цих же таблеток ($SD^2(UM)$) – див. Табл. 13.П-8.

Таким образом, в цьому випадку в рівнянні (13.П-1) маємо:

$$y = SD^2(UC); \quad x = SD^2(UM).$$

- 3) Залежність відношення кількісного вмісту в одній таблетці до маси тієї же таблетки ($100 \times UC_{ij} / UM_{ij}$) від середньої маси тих же таблеток ($100 \times SM(UM)$) – див. Табл. 13.П-8. Дана кореляція дозволяє підтвердити (або не підтвердити), що між кількісним вмістом в одній таблетці і її масою існує лінійна залежність.

По аналогії з п. 1, доцільно використовувати не самі величини, а їх відхилення від 100%.

Таким чином, в цьому випадку в рівнянні (13.П-1) маємо:

$$y = \overline{100 \times UC_{ij} / UM_{ij}} - 100; \quad x = 100 \times SM(UM) - 100$$

По аналогії з п. 2, нас цікавить залежність дисперсій величин ($100 \times UC_{ij} / UM_{ij}$) від дисперсії маси тих же таблеток, тобто в цьому випадку в рівнянні (13.П-1) маємо:

$$y = SD^2(\overline{100 \times UC_{ij} / UM_{ij}}); \quad x = SD^2(UM).$$

Таблиця 13.П-10

Таблиця дисперсійного аналізу для однорідності вмісту

	Головний ефект				
	ν	SS	MS	F	$P(>F)\%$
Міжгрупова (bg) варіація ($effect$)	8	166.69	20.84	8.02	0.00 < 1.0 %
Внутрішньогрупова (wg) варіація ($error$)	81	210.56	2.60		
В цілому (tot)	89	377.25	4.24		
Коефіцієнт детермінації η^2 (%)					44.2

Таблиця 13.П-11

Таблиця дисперсійного аналізу для величини $100 \times UC_{ij} / UM_{ij} \%$

	Головний ефект				
	ν	SS	MS	F	$P(>F)\%$
Міжгрупова (bg) варіація ($effect$)	8	33.34	4.17	19.53	0.00 < 1.0 %
Внутрішньогрупова (wg) варіація ($error$)	81	17.28	0.21		
В цілому (tot)	89	50.63	0.57		
Коефіцієнт детермінації η^2 (%)					65.9

Ця залежність дозволяє підтвердити (або не підтвердити), що однорідність вмісту визначається однорідністю маси.

Оцінка значущості отриманих метрологічних величин проводиться за t -критерієм (ймовірність того, що величини b і a дорівнюють нулю) і за F -критерієм (ймовірність того, що загальний індекс кореляції R_c дорівнює нулю).

Значущість коефіцієнтів a і b

Для оцінки значущості відмінності величин b і a від нуля розраховують t -значення за співвідношенням:

$$\begin{aligned} t_b &= b / s_b, \\ t_a &= a / s_a. \end{aligned} \quad (13.П-3)$$

На основі цих t -значень розраховують ймовірність $P(>t_b, \nu) \%$ і $P(>t_a, \nu) \%$ того, що величини a або b дорівнюють нулю. Число ступенів свободи для t -критерію дорівнює $\nu = k - 2 = 9 - 2 = 7$.

Ймовірність менше 5.0 % свідчить про значущу, а ймовірність менше 1.0 % - про високозначущу відмінність величин a або b від нуля.

Значущість загального індексу кореляції R_c

Для оцінки значущості загального індексу кореляції R_c від нуля використаємо співвідношення (13.14) для зв'язку індексу кореляції R_c з F -критерієм, яке представимо в такому вигляді:

$$\begin{aligned} F_{s(SM)/s_0} &= \frac{1}{1 - R_c^2}. \\ P(>F, \nu_y, \nu_{s_0}) \% &< \alpha \quad (13.П-4) \\ \nu_y &= k - 1. \\ \nu_{s_0} &= k - 2. \end{aligned}$$

$P(>F, \nu_y, \nu_{s_0})$ представляє собою ймовірність того, що загальний індекс кореляції R_c дорівнює нулю. Критичне значення цієї ймовірності задається рівнем значущості α . Вважається, що якщо $P(>F, \nu_y, \nu_{s_0}) < \alpha = 5.0 \%$, то R_c значуще, а при $P(>F, \nu_y, \nu_{s_0}) < \alpha = 1.0 \%$ - високозначуще відрізняється від нуля.

В нашому випадку $k = 9$, і числа ступенів свободи дорівнює $\nu_y = k - 1 = 9 - 1 = 8$, $\nu_{s_0} = k - 2 = 7$.

Продемонструємо застосування регресійного аналізу на прикладі залежності кількісного вмісту від середньої маси таблеток. Інші залежності отримують за тією ж схемою.

13.10.3.4.2. Залежність кількісного вмісту від середньої маси таблеток

Як же згадувалося вище, в цьому випадку в рівнянні (13.П-1) маємо:

$$y = SM(UC) - 100; \quad x = SM(UM) - 100.$$

Проводячи обробку за даним рівнянням значень $SM_i(UC)$ і $SM_i(UM)$, які представлені в Табл. 13.П-8, за допомогою ЛМНК за схемою, описаною в розділі

12.9.1, отримуємо наступні метрологічні характеристики цієї прямої ($k = 9$, $\nu = k - 2 = 7$):

$$b = 1.40, \quad s_b = 0.22, \quad t_b = 1.40 / 0.22 = 6.29, \\ P(>t_b = 6.29, \nu = 7) = 0.02 < 1.0 \%$$

$$a = 2.50, \quad s_a = 0.32, \quad t_a = 2.50 / 0.32 = 7.90, \\ P(>t_a = 7.90, \nu = 7) = 0.00 < 1.0 \%$$

$$s_0 = 0.59, \quad R_c = 0.91.$$

Як видно, коефіцієнти b і a високозначуще відрізняються від нуля – ймовірність того, що вони дорівнюють нулю, менше 1.0 %

Перевірка значущості коефіцієнта кореляції

Розрахуємо F -значення і відповідну йому ймовірність:

$$F\text{-значення} = 1(1 - 0.91^2) = 5.82.$$

$$P(>F = 5.82, \nu_y = 9 - 1, \nu_{s_0} = 9 - 2) = 1.57 \% < 5.0 \%$$

Як видно, ймовірність того, що коефіцієнт кореляції дорівнює нулю, менше 5.0 %, тобто він значуще відрізняється від нуля.

Таким чином, можна зробити висновок, що кількісний вміст в одній таблетці статистично значуще лінійно залежить від маси тієї ж таблетки. Одночасно коефіцієнти b і a цієї ж лінійної залежності (13.П-2) високозначуще відрізняються від нуля.

Розрахунки для других регресій проводять за тією ж схемою. Результати регресійного аналізу підсумовані в Табл. 13. П-12.

13.10.3.4.3. Обговорення результатів регресійного аналізу

Як видно з Табл. 13.П-12, спостерігається значуща лінійна залежність (1) між кількісним вмістом і масою таблеток ($R_c = 0.91$). Ймовірність того, він дорівнює нулю, складає $1.57 < 5.0 \%$, тобто R_c значуще відрізняється від нуля. Одночасно коефіцієнти b і a високозначуще відрізняються від нуля: ймовірність того, що вони дорівнюють нулю, менше 1.0 %. Висока значущість вільного члена a свідчить про те, що кількісний вміст визначається не тільки масою таблеток, але й іншими факторами.

Дисперсії SD^2 характеризують однорідність результатів. Як видно, спостерігається високозначуща лінійна залежність (2) Табл. 13.П.12 між однорідністю вмісту і однорідністю маси ($R_c = 0.96$). Ймовірність того, що R_c дорівнює нулю, складає $0.01 < 1.0 \%$, тобто він високозначуще відрізняється від нуля. Це підтверджується і високою значущістю кута нахилу b . В той же час, вільний член a є незначущим. Це свідчить про те, що однорідність вмісту в даному випадку практично повністю визначається однорідністю маси. Цей важливий висновок дозволяє контролювати однорідність вмісту за значно простішим випробуванням на однорідність маси.

Метрологічні характеристики рівняння (3) Табл. 13.П-12 свідчать про відсутність лінійної за-

лежності для відношення кількісного вмісту до маси таблеток – $R_c = 0.42$ і незначуще відрізняється від нуля (ймовірність того, що він дорівнює нулю, складає 40.7 %). Це підтверджує те, що залежність (1) вмісту від маси є лінійною. Незначущим є також кут нахилу b . В той же час, вільний член $a = 2,51$ є високозначущим і співпадає з вільним членом рівняння (1) Табл. 13.П.12 ($a = 2.5$). Це підтверджує висновок пункту 1 про те, що кількісний вміст визначається не тільки масою таблеток, але й другими факторами.

Всі параметри рівняння (4) Табл. 13.П-12 незначуще відрізняються від нуля. Це підтверджує висновок пункту 2 про те, що однорідність вмісту практично повністю визначається однорідністю маси – оскільки дисперсія їх відношення не залежить від дисперсії маси.

Таблиця 13.П-12

Результати регресійного аналізу експериментальних даних для прямої $y = b \times x + a$.

Величина	Значення	SD	t	F	$P(>t, \nu) \%$	Висновок
(1) $y = SM(UC) - 100; x = SM(UM) - 100$						
b	1.40	0.22	6.29		0.02	< 1.0% високозначущий
a	2.50	0.32	7.90		0.00	< 1.0% високозначущий
R_c	0.91	0.59*		5.82	1.57**	< 5.0 % значущий
(2) $y = SD^2(UC); x = SD^2(UM)$						
b	1.31	0.13	10.17		0.00	< 1.0% високозначущий
a	-0.060	0.29	0.19		43.0	> 5.0 % незначущий
R_c	0.96	0.39*		13.8	0.01**	< 1.0% високозначущий
(3) $y = 100 \times UC / UM - 100; x = SM(UM) - 100$						
b	0.36	0.22	1.64		7.27	> 5.0 %, незначущий
a	2.51	0.31	8.01		0.00	< 1.0% високозначущий
R_c	0.42	0.58*		1.21	40.7**	> 5.0 %, незначущий
(4) $y = SD^2(100 \times UC / UM); x = SD^2(UM)$						
b	0.069	0.046	1.51		8.75	> 5.0 %, незначущий
a	0.073	0.104	0.70		25.35	> 5.0 %, незначущий
R_c	0.37	0.14*		1.160	42.9**	> 5.0 %, незначущий

* S_0

** $P(>F, \nu_y = 8, \nu_{s_0} = 7)$