

**МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ  
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ВНУТРІШНІХ СПРАВ**

**Факультет № 6  
Кафедра соціології та психології**

**ТЕКСТ ЛЕКЦІЇ**

з навчальної дисципліни  
**«Комп'ютерні методи практичної психології»**  
обов'язкових компонент освітньої програми  
першого (бакалаврського) рівня вищої освіти

**053 Психологія (практична психологія)**

**Тема №8. Аналіз психологічних даних в SPSS: дисперсійний аналіз**

**Харків 2022**

**ЗАТВЕРДЖЕНО**

Науково-методичною радою  
Харківського національного  
університету внутрішніх справ  
Протокол від 30.08.2022 р. №8

**СХВАЛЕНО**

Вченою радою факультету №6  
Протокол від 26.08.2022 р. №6

**ПОГОДЖЕНО**

Секцією Науково-методичної  
ради ХНУВС з гуманітарних та  
соціально-економічних дисциплін  
Протокол від 29.08.2022 р. №8

Розглянуто на засіданні кафедри соціології та психології  
Протокол від 17.08.2022 р. №7

**Розробник:**

Професор кафедри соціології та психології факультету №6,  
д-р соціол. н., професор Нечитайло Ірина Сергіївна

**Рецензенти:**

1. Керівник психологічної служби Харківського гуманітарного університету «Народна українська академія», доцент кафедри соціології та гуманітарних дисциплін, к. психол. н., Гога Н. П.;
2. Доцент кафедри соціології та психології факультету №6, к. психол. н., Філоненко В. М.

## ТЕМА №8. АНАЛІЗ ПСИХОЛОГІЧНИХ ДАНИХ В SPSS: ДИСПЕРСІЙНИЙ АНАЛІЗ

### План

- 8.1.** Поняття дисперсійного аналізу.
- 8.2.** Алгоритми обчислення, які стосуються дисперсійного аналізу.

### Рекомендована література

#### *Основна*

1. Максименко С. Д., Носенко Е. Л. Експериментальна психологія : підручник Київ : Центр учб. літ., 2017. 359 с.
2. Нечитайло И. С., Бирюкова М. В. Математические методы в социологии : Учебник для студентов высш. учеб. заведений / Нар. укр. акад., [каф. социологии]. Харьков : Изд-во НУА, 2012. 243 с.
3. Осенкив Г. С. Психология массового поведения : учеб.-практ. пособие Киев : Сварог, 2017. 158 с.
4. Татьянчиков А. О. Математичні методи в психології: навчально-методичні рекомендації (в допомогу до самостійної роботи для здобувачів вищої освіти ступеня бакалавра факультету психології, політології та соціології) ; кафедра психології НУ «Одеська юридична академія». Одеса : Фенікс, 2021. 48 с.
5. Телейко А. Б., Чорней Р. К. Математико-статистичні методи в соціології та психології : Навч. Посібник. Київ : МАУП, 2007. 424 с.

#### *Допоміжна*

6. Застосування математично-статистичних методів аналізу у психологічних вимірюваннях [Електронний ресурс]. Режим доступу: <http://surl.li/aghly> . Дата звернення: 01.09.2021.
7. Катаєв Є. С. Використання статистичних методів обробки даних у дослідженнях “я-концепції” особистості. Вісник Національного університету оборони України. 2012. №2 (27) /2012. С. 171-176.
8. Салюк М. А. Статистична обробка даних експериментального дослідження. Методичний посібник з курсу «Експериментальна психологія» / за ред. Е.Л. Носенко. Дніпропетровськ: Інновація, 2010. 26 с.
9. Старушенко Г. А. Статистична обробка даних в системі публічного управління : навч. посіб. Дніпро : ГРАНІ, 2018. 144 с.
10. Foster G., Lane D., Scott D., Hebl M. and other. An Introduction to Psychological Statistics. University of Missouri, St. Louis. 2018. 271 p.

## Текст лекції

### 8.1. Поняття дисперсійного аналізу

*Дисперсія* (англ. variance) – це міра розсіювання значень випадкової величини відносно середнього значення розподілу. Більші значення дисперсії свідчать про більші відхилення значень випадкової величини від центру розподілу.

У простому розумінні, дисперсія дозволяє виміряти наскільки далеко випадкові значення розподілені від їх середнього значення. Дисперсія відіграє важливу роль в статистиці, в якій вона використовується в таких напрямках як описова статистика, перевірка статистичних гіпотез, і Метод Монте-Карло. Дисперсія дорівнює квадрату стандартного відхилення, що є другим центральним моментом розподілу, і коваріації випадкової величини із самою собою.

*Дисперсія випадкової величини* – це один з параметрів розподілу ймовірностей, середньоквадратичне відхилення від середнього значення. Інакше кажучи, це математичне сподівання квадрату відхилення цієї змінної від її очікуваного значення (її математичного сподівання). Отже дисперсія є вимірюванням величини розпорошеності значень цієї змінної, беручи до уваги всі її значення і їхні ймовірності або ваги.

*Дисперсійний аналіз.* Основною метою дисперсійного аналізу, фундаментальна концепція якого була запропонована Фішером у 1920 р., є дослідження значущості відмінності між середніми декількох груп даних або змінних. Якщо порівнюються середні двох груп, дисперсійний аналіз дасть той же результат, що і звичайний  $t$ -критерій для незалежних або залежних вибірок. Проте використання дисперсійного аналізу має переваги особливо для малих вибірок. У дисперсійному аналізі перевірка статистичної значущості відмінності між середніми декількох груп здійснюється на основі вибірових дисперсій. Ця перевірка проводиться за допомогою розбиття загальної дисперсії (варіації) на частини, одна з яких обумовлена випадковою помилкою (тобто внутрішньогруповою мінливістю), а друга пов'язана з відмінністю середніх значень. Якщо ця відмінність значуща, нульова гіпотеза щодо існування відмінності між середніми значеннями відкидається на певному рівні значущості.

*Дисперсійний однофакторний аналіз* використовується у дослідженнях зміни результативної ознаки під впливом зміни умов або градацій фактору. Суть математичних перетворень дисперсійного методу полягає в тому, щоб зіставити дисперсії за факторами із дисперсією усіх значень, отриманих в експерименті. Однофакторний аналіз вимагає не менше трьох градацій фактору і не менше двох випробувань у кожній градації. При проведенні дисперсійного аналізу необхідно перевірити нормальність розподілу досліджуваної випадкової величини і відсутність відмінності дисперсій сукупностей. Це можна виконати методами перевірки статистичних гіпотез

Однофакторний дисперсійний аналіз в SPSS реалізується за допомогою

команди One-way ANOVA. Команди підміни General Linear Model (Загальні лінійні моделі), також частково дозволяють проводити подібний аналіз, однак їх можливості є більш обмеженими, ніж у команди One-way ANOVA (Однофакторний дисперсійний аналіз).

Дисперсійний аналіз – це процедура порівняння середніх значень вибірок, на підставі якої можна зробити висновок про співвідношення середніх значень генеральних сукупностей. Найближчим і більш простим аналогом ANOVA є t-критерій. На відміну від t-критерію дисперсійний аналіз призначений для порівняння не двох, а кількох вибірок. Слово «дисперсійний» у назві вказує на те, що в процесі аналізу зіставляються компоненти дисперсії досліджуваної змінної.

Загальна мінливість змінної розкладається на дві складові: міжгрупову (факторну), обумовлену різницею груп (середніх значень), і внутрішньогрупову (помилки), обумовлену випадковими (врахованими) причинами. Чим більшою є частка від ділення міжгрупової і внутрішньогрупової мінливості (F-відношення) тим більшою мірою відрізняються середні значення порівнюваних вибірок і тим вищою є статистична значущість цієї відмінності.

Отже, сама назва процедури вказує на те, що висновок про відмінність середніх значень робиться на основі аналізу компонентів дисперсії. Що ж означає термін «Однофакторний»? Застосовуючи команду One-way ANOVA (Однофакторний дисперсійний аналіз), ви побачите, що виконання цієї команди передбачає вказування єдиної залежної змінної (при цьому вона обов'язково повинна бути кількісного, а точніше, метричного типу) і єдиної незалежної змінної (завжди номінальної (номінативної), що має декілька градацій). Різні моделі дисперсійного аналізу, описані у двох наступних главах, допускають наявність кількох незалежних змінних.

При однофакторному дисперсійному аналізі порівнюються середні значення кожної вибірки одне з одним і обчислюється загальний рівень значущості відмінностей. Зверніть увагу, що висновок за результатами ANOVA стосується загальної відмінності всіх порівнюваних середніх без конкретизації того, які вибірки розрізняються, а які – ні. Для ідентифікації пар вибірок, що відрізняються одна від одної середніми значеннями, використовуються методи парних порівнянь постфактум (post hoc), а для більш складних зіставлень – метод контрастів (contrasts).

## **8.2. Алгоритми обчислення, які стосуються дисперсійного аналізу**

### *Покрокові алгоритми обчислень*

Для проведення однофакторного дисперсійного аналізу слід використовувати робочий файл. У ролі залежної змінної виступить мінна «тест 1», а незалежна змінна «клас» розділить об'єкти на 3 вибірки, середні значення яких ми будемо порівнювати, але спочатку необхідно виконати три підготовчих кроки:

**Крок 1.** Створіть новий файл даних або підготуйте існуючий.

**Крок 2.** Запустіть програму SPSS за допомогою значка на робочому столі або команди Пуск > Програми > SPSS for Windows > SPSS 11.5 for Windows (Start > 1(1 Programs > SPSS for Windows > SPSS 11.5 for Windows) головного меню Windows.

Після запуску програми діалоговому вікні SPSS for Windows клацніть на кнопці Cancel (Скасувати). Після виконання цього кроку на екрані з'явиться вікно редактора даних SPSS.

**Крок 3.** Відкрийте файл даних, з якими ви маєте намір працювати. Якщо він розташований в поточній папці, то виконаєте наступні дії:

1. Виберіть у меню File (Файл) команду Open > Data (Відкриття > Дані) або клацніть на кнопці > Open File (відкрити файл) панелі інструментів.
2. У діалоговому вікні, двічі клацніть на імені файлу або введіть з клавіатури і клацніть на кнопці OK.

Незалежно від того, відкрита програма SPSS тільки-но або якісь процедури вже виконувалися, у верхній частині головного вікна повинна бути присутнім рядок меню (вона показана нижче). Поки рядок меню присутня на екрані, доступні всі команди аналізу даних. При цьому вікно з даними бачити не обов'язково.

При роботі з таблицями чи під час редагування діаграм деякі пункти меню можуть зникати або змінюватися. Щоб повернутися до головного вікна і стандартної панелі меню, клацніть на кнопці згортання або відновлення поточного вікна.

Після завершення кроку 3 на екрані має з'явитися вікно редактора даних з рядком меню.

**Крок 4.** В меню Analyze (Аналіз) виберіть команду Compare Means > One-Way ANOVA (Порівняння середніх > Однофакторний дисперсійний аналіз). На екрані з'явиться діалогове вікно One-Way ANOVA (Однофакторний > дисперсійний аналіз).

Структура діалогового вікна One-Way ANOVA (Однофакторний дисперсійний аналіз) цілком типова для більшості діалогових вікон SPSS. Зліва ми бачимо список змінних поточного файлу даних. У нижній частині вікна розташовані три кнопки: Options (Параметри), Post Hoc (Постфактум) і Contrasts (Контрасти), які ми будемо використовувати при обробці. Список Dependent List (Залежні змінні) призначений для задання однієї або кількох залежних змінних (у нашому прикладі буде використовуватися єдина залежна змінна «тест 1»). Залежні змінні повинні бути метричного типу. Якщо в списку вказано кілька залежних змінних, то SPSS виконає ANOVA для кожної з них. Під списком Dependent List (Залежні змінні) знаходиться поле Factor (Фактор), в якому потрібно вказати єдину незалежну змінну, що має кілька градацій (в нашому випадку це буде «хобі»). Таким чином ми порівнюємо результати першого тесту («рахунок в умі») для трьох груп учнів, що розрізняються позашкільними захопленнями.

Іноді виникає необхідність порівнювати не всі вибірки, відповідні градаціях фактору (незалежної змінної), а лише частина з них. У цьому випадку перед проведенням аналізу необхідно звернутися до команди Select

Cases (Вибір об'єктів) для вибору необхідних для аналізу градацій фактору.

На кроці 5 будемо порівнювати між собою середні значення змінної «тест 1» для кожної з вибірок за рівнями змінної хобі.

**Крок 5.** Після виконання кроку 4 має бути відкрито діалогове вікно One-Way ANOVA (однофакторний дисперсійний аналіз).

При необхідності повторіть крок 4 і виконайте наступні дії:

1. Клацніть спочатку на змінну «тест 1», щоб виділити її, а потім – на верхній кнопці зі стрілкою, щоб перемістити змінну в список Dependent List (Залежні змінні).

2. Клацніть спочатку на змінну «хобі», щоб виділити її, а потім – на нижній кнопці зі стрілкою, щоб перемістити змінну в поле Factor (Фактор).

3. Клацніть на кнопці ОК, щоб відкрити вікно виведення. Існує дві додаткові дії, які бажано виконувати в процесі аналізу. Наведена послідовність інструкцій дозволяє отримати результати порівняння середніх значень вибірок, однак ані самі середні значення, ані результати парного порівняння вибірок у виведених даних відображені не будуть.

Перша проблема вирішується за допомогою кнопки Options (Параметри). Діалогове вікно One-Way ANOVA: Options (Однофакторний дисперсійний аналіз: Параметри), що з'являється після клацання на ній. Установка прапорця Descriptive (Описові статистики) призведе до включення у дані о виводяться, усіх середніх значень, стандартних відхилень, стандартних помилок, меж довірчих інтервалів у 95%, а також мінімумів і максимумів вибірок. Прапорець Homogeneity of variance test (Критерій однорідності дисперсії) дозволяє вивести інформацію про ступінь придатності даних до дисперсійного аналізу, а з допомогою прапорця Means plot (Графік середніх) можна побудувати діаграму, на якій будуть зображені середні значення для кожної вибірки. Група перемикачів Missing Values (Пропущені значення) дозволяє вибрати спосіб обробки відсутніх значень.

*Парні порівняння.* Нерідко нас можуть зацікавити результати парних порівнянь градацій незалежної змінної, і для цієї мети в діалоговому вікні One-Way ANOVA (Однофакторний дисперсійний аналіз) передбачена спеціальна кнопка Post Hoc (Постфактум). Слово «постфактум» означає, що ця процедура проводиться після встановлення статистично достовірного результату однофакторного дисперсійного аналізу. Якщо результати ANOVA виявилися статистично недостовірними, застосування процедури парних порівнянь некоректно. При клацанні на кнопці Post Hoc (Постфактум) відкривається діалогове вікно One-Way ANOVA: Post Hoc Multiple Comparisons (Однофакторний дисперсійний аналіз: Множинні порівняння постфактум). Це діалогове вікно за допомогою прапорців дозволяє задати 14 критеріїв для вибірок з однаковою дисперсією і 4 критерію для вибірок з різною дисперсією. Число пропонованих критеріїв настільки велике, що навіть керівництво користувача програми SPSS обсягом близько 3000 сторінок виявилось не в змозі вмістити опису їх усіх.

Більшість із зазначених тестів використовуються дуже рідко, тому нижче наведені описи лише декількох найбільш популярних.

LSD (Least Significant Difference – Найменша значуща різниця) – цей тест являє собою не що інше, як сукупність t-критеріїв для всіх можливих пар градацій фактору. Критерій найменшою значущою різниці є одним з найбільш «ліберальних», оскільки найбільш схильний до помилок. Наприклад, якщо незалежна змінна має 5 рівнів, то буде проведено 10 порівнянь. При рівнях значущості кожного з порівнянь, рівних 0,05, існує ймовірність майже у 40% того, що хоча б 1 з тестів показав значимий результат випадково.

Bonferroni (Бонферроні) – критерій Бонферроні схожий з критерієм найменшої значимої різниці, однак позбавлений недоліку, пов'язаного з повторними перевірками: у ньому рівень значущості ділиться на число порівнянь. Таким чином, критерій Бонферроні є більш «консервативним».

Scheffe (Шеффе) – критерій Шеффе ще більш «консервативний», ніж критерій Бонферроні, використовує f-критерій замість t-критерію.

t-Tukey (Тьюки) – критерій Тьюки використовує статистику Стюдента (Student) для визначення відмінностей між групами. Цей критерій часто застосовується у випадках, коли досліджуваний фактор має велику кількість рівнів.

Найбільш консервативними із запропонованих є критерії Шеффе і Бонферроні. Часто використовується також критерій Тьюки, званий ще «критерієм справжньої значущості» (Honestly Significant Difference, HSD). HSD являє собою найменшу величину різниці середніх значень вибірок, яку можна вважати значущою. Наприклад, якщо HSD – 2,5, а для двох вибірок отримані величини середніх значень 3,7 і 6,3, то різниця між ними, що дорівнює 2,6, згідно Тьюки, є значущою, оскільки вона перевищує величину HSD.

При використанні критерію t-Tukey програма SPSS також включає в висновок додаткову статистичну інформацію. Необхідно відзначити, що всі наведені вище критерії (а насправді більшість критеріїв парних порівнянь) застосовуються в припущенні, що дисперсії всіх осередків рівні. Виняток становлять 4 критерію, прапорці для яких виділені в окрему групу в нижній частині діалогового вікна One-Way ANOVA: Post Hoc Multiple Comparisons (Однофакторний дисперсійний аналіз: Множинні порівняння постфактум), – вони застосовуються у випадках, коли дисперсії осередків різні.

У наступному прикладі ми проведемо однофакторний дисперсійний аналіз, а у результати, що виводяться, включимо описові статистики і критерій однорідності дисперсії. Для парних порівнянь скористаємося критерієм Шеффе.

Як і в попередньому прикладі, залежною змінною буде «тест 1», а незалежною – «хобі».

**Крок 5а.** Після виконання кроку 4 має відкритися діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз). Якщо ви вже встигли попрацювати з цим вікном, очистіть його клацанням на кнопці Reset (Скидання) і виконайте такі дії:

1. Клацніть спочатку на змінну «тест 1», щоб виділити її, а потім – на кнопці вгорі зі стрілками, щоб перемістити змінну в список Dependent List



(Залежні змінні);

2. Клацніть спочатку на змінну «хобі», щоб виділити її, а потім – на нижній кнопці зі стрілкою, щоб перемістити змінну в поле Factor (Фактор);

3. Клацніть на кнопці Options (Параметри), щоб відкрити діалогове вікно One-Way ANOVA: Options (Однофакторний дисперсійний аналіз: Параметри);

4. Установіть прапорці Descriptive (Описові статистики) і Homogeneity of variance test (Критерій однорідності дисперсії), а потім клацніть на кнопці Continue (Продовжити), щоб повернутися в діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз);

5. Клацніть на кнопці Post Hoc (Постфактум), щоб відкрити діалогове вікно One-Way ANOVA: Post Hoc Multiple Comparisons (Однофакторний дисперсійний аналіз: Множинні порівняння постфактум);

6. Установіть прапорець Sheffe (Шеффе) і клацніть на кнопці Continue (Продовжити), щоб повернутися в діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз);

7. Клацніть на кнопці ОК, щоб відкрити вікно виведення.

Остання з трьох кнопок у вікні One-Way ANOVA (Однофакторний дисперсійний аналіз) має назву Contrasts (Контрасти). Вона призначена для виклику діалогового вікна One-Way ANOVA: Contrasts (Однофакторний дисперсійний аналіз: Контрасти). Це вікно дозволяє здійснювати різні порівняння вибірок по градаціям незалежної змінної. Так, ви можете порівнювати одну градацію з іншого, одну градацію з усіма іншими або розбити всі градації на 2 групи і потім порівняти їх між собою.

Порівняння зводиться до застосування модифікованого варіанта t-критерію.

Зазначимо, що на відміну від процедури парного порівняння постфактум застосування методу контрастів не потребує попереднього отримання статистично достовірного результату аналізу ANOVA.

У нашому прикладі незалежна змінна «хобі» має 3 градації: 1 – спорт, 2 – комп'ютер, 3 – мистецтво. Для завдання контрасту призначені поле і список Coefficients (Коефіцієнти). Заповнення списку відбувається наступним чином. З кожною градацією фактору ви повинні зіставити число, що визначає його роль в контрасті: від'ємне число відповідає одній групі, додатне число – іншій групі, а нуль означає, що градація у порівняннях не задіяна. При цьому абсолютні величини коефіцієнтів не важливі, але останні повинні вводитися в порядку прямування градацій і в сумі давати нульове значення. Наприклад, якщо вам необхідно порівняти тих учнів, які захоплюються спортом, з іншими учнями, ви можете закодувати градації послідовністю: -2, 1 і 1, а якщо ви хочете порівняти тих учнів, які захоплюються спортом, з тими, хто захоплюється мистецтвом, то необхідно задати послідовність: 1, 0 і -1. Не забувайте, що сума всіх коефіцієнтів обов'язково повинна дорівнювати 0.

Розміщення чисел у списку здійснюється шляхом введення значень у полі праворуч від назви списку і клацання на кнопці Add (Додати). Якщо потрібно створити кілька «контрастів», тобто розбити на групи, необхідно

клацнути на кнопці Next (Наступний) праворуч від мітки Contrast 1 of 1 (Контраст 1 з 1) і повторіть процедуру.

Наступний приклад ілюструє застосування контрастів для порівняння учнів, які захоплюються комп'ютером, з тими, хто має інші захоплення, а також для порівняння тих, хто захоплюється спортом з іншими учнями.

Після виконання кроку 5 програма автоматично активізує вікно виведення. Для перегляду результатів, при необхідності, можна скористатися вертикальною і горизонтальною смугами прокрутки. Зверніть увагу на стандартний рядок меню у верхній частині вікна виводу: його присутність дозволяє виконувати будь-які статистичні операції, не перемикаючись назад у вікно редактора даних.

**Крок 5б.** Після виконання кроку 4 має відкритися діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз).

Якщо ви вже встигли попрацювати з цим вікном, очистіть його клацанням на кнопці Reset (Скидання) і виконайте такі дії:

1. Клацніть спочатку на змінну «тест 1», щоб виділити, а потім – на кнопці вгорі зі стрілками, щоб перемістити змінну в список Dependent List (Залежні змінні);
  2. Клацніть спочатку на змінну хобі, щоб виділити її, а потім – на нижній кнопці зі стрілкою, щоб перемістити змінну в поле Factor (Фактор);
  3. Клацніть на кнопці Options (Параметри), щоб відкрити діалогове вікно One-Way ANOVA: Options (Однофакторний дисперсійний аналіз: Параметри);
  4. Установіть прапорці Descriptive (Описові статистики) і Homogeneity of variance test (Критерій однорідності дисперсії), а потім клацніть на кнопці Continue (Продовжити), щоб повернутися в діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз);
  5. Клацніть на кнопці Post Hoc (Постфактум), щоб відкрити діалогове вікно One-Way ANOVA: Post Hoc Multiple Comparisons (Однофакторний дисперсійний аналіз: Множинні порівняння постфактум);
  6. Установіть прапорець LSD (Найменша значуща різниця) і клацніть на кнопці Continue (Продовжити), щоб повернутися в діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз);
  7. Клацніть на кнопці Contrasts (Контрасти), щоб відкрити діалогове вікно One-Way ANOVA: Contrasts (Однофакторний дисперсійний аналіз: Контрасти);
  8. Натисніть клавішу Tab, щоб перевести курсор в полі Coefficients (Коефіцієнти), введіть 1 і клацніть на кнопці Add (Додати), задавши перше число в списку;
  9. Повторіть попередній крок спочатку для чисел 2 і 1, потім, клацнувши на кнопці Next (Наступний), – для чисел -2, 1, 1, після чого клацніть на кнопці Continue (Продовжити), щоб повернутися в діалогове вікно One-Way ANOVA (Однофакторний дисперсійний аналіз);
  10. Клацніть на кнопці ОК, щоб відкрити вікно виведення.
- Багатофакторний дисперсійний аналіз відрізняється від однофакторного

появою однієї загальної проблеми – проблеми взаємодії факторів. Рішення цієї проблеми принципово не залежить від кількості факторів. Зі зростанням числа факторів істотно наростає лише складність інтерпретації взаємодій. Тому ми детально розглянемо тільки найпоширеніший варіант дисперсійного аналізу – аналіз з двома факторами, припускаючи, що отримані вміння ви зможете поширити і на аналіз з великим числом факторів.

В рамках цієї теми виконаємо двухфакторний дисперсійний аналіз і аналіз з урахуванням впливу коваріати.

*Дисперсійний аналіз з двома факторами.* Як було вже сказано, дисперсійний аналіз визначає статистичну достовірність відмінності між вибірками шляхом порівняння їх середніх значень. Щоб отримати уявлення про двофакторний аналіз, спочатку трохи узагальнимо знання про однофакторний дисперсійний аналіз.

Ми порівнювали три класи (змінна «клас») за рівнем вираженості змінної «тест 1». «Однофакторність» аналізу полягала у тому, що розподіл на групи проводився за градаціями однієї незалежної змінної («клас»).

Для однофакторного дисперсійного аналізу в SPSS існує спеціальна спрощена команда One-Way ANOVA (однофакторний дисперсійний аналіз). Команди підміню General Linear Model (Загальні лінійні моделі) дозволяють виконувати однофакторний, двофакторний, трифакторний тощо аналіз, однак виконання останніх двох операцій є трохи складнішим.

Будемо використовувати звичний для нас робочий файл. В якості залежної змінної візьмемо змінну «тест 2», а роль незалежних будуть відігравати змінні «стать» і «хобі». Ми спробуємо визначити ступінь впливу змінних «стать» і «хобі» на розподіл значень змінної «тест 2». Така схема аналізу може бути позначена як ANOVA 2x3 («стать» x «хобі»). Дослідження дозволить отримати відповіді на такі запитання:

- 1) Чи існує головний ефект фактору «стать», тобто чи існує значуща відмінність оцінок юнаками та дівчатами, і якщо так, то якою є ступінь відмінності?
- 2) Чи існує головний ефект фактору «хобі», тобто чи існує значуща відмінність оцінок між трьома групами (які означають різні види хобі), і якщо так, то якою є ступінь відмінності?
- 3) Чи існує взаємодія змінних «стать» і «хобі», тобто чи впливає одна на іншу?

Таким чином, двофакторний дисперсійний аналіз дозволяє перевірити три гіпотези: дві – про головні ефекти і одну – про взаємодію факторів.

Відповіді на перші два питання можна було б отримати, застосувавши двічі однофакторний дисперсійний аналіз. Специфіка багатофакторного аналізу проявляється в змісті третього питання, що стосується взаємодії факторів.

Взаємодія двох чинників означає, що вплив одного з них проявляється по-різному на різних рівнях іншого фактору. Вельми корисним для інтерпретації взаємодій є побудова діаграми середніх значень для кожної комірочки таблиці спряженості незалежних змінних. На діаграмах представлені

всі значення незалежно від ступеня їх значущості, що дозволяє більш чітко уявляти, чому вплив одних змінних виявляється значущим, а інших – ні.

SPSS може включити до висновку середні значення вибірок, що відповідають усім можливим поєднанням градацій факторів (в даному випадку  $2 \times 3 = 6$ ), а також обчислити F-величини і відповідні  $p$ -рівні. За цими характеристиками ми зможемо судити про ступінь впливу кожної з незалежних змінних на розподіл залежної змінної, а також про взаємодію незалежних змінних.

*Дисперсійний аналіз з трьома і більше факторами.* Як вже зазначалося, проведення дисперсійного аналізу з трьома і більше факторами принципово не відрізняється від процедури двофакторного аналізу. При кількості факторів більше двох зростає лише складність інтерпретації взаємодій факторів. Ця складність зумовлена появою великого числа взаємодій, деякі з яких важко піддаються інтерпретації. Розглянемо цю проблему на прикладі. Припустимо, вивчається вплив на змінну «тест 2» трьох факторів: «стать», «хобі» та «клас». Така схема аналізу може бути позначена як ANOVA  $2 \times 3 \times 3$  («стать»  $\times$  «хобі»  $\times$  «клас»). Застосування трифакторного дисперсійного аналізу дозволило б отримати відповіді на наступні питання:

- 1) Чи існує головний ефект фактору «стать», тобто чи існує значуща відмінність оцінок для юнаків і дівчат, і який ступінь цієї відмінності?
- 2) Чи існує головний ефект фактору «хобі», тобто чи існує значуща відмінність оцінок між трьома групами, і який ступінь цієї відмінності?
- 3) Чи існує головний ефект фактору «клас», тобто чи існує значуща відмінність оцінок між класами, і який ступінь цієї відмінності?
- 4) Чи існує взаємодія змінних «стать» і «хобі», тобто чи залежить вплив однієї з цих змінних від градацій іншої?
- 5) Чи існує взаємодія змінних «стать» і «клас», тобто чи залежить вплив однієї з цих змінних від градацій іншої?
- 6) Чи існує взаємодія змінних «клас» і «хобі», тобто чи залежить вплив однієї з цих змінних від градацій іншої?
- 7) Чи існує взаємодія змінних «стать», «хобі» та «клас», тобто чи залежить взаємодія двох з цих змінних від градацій третьої?

Таким чином, трифакторний дисперсійний аналіз передбачає перевірку вже семи гіпотез. З них три гіпотези стосуються взаємодії першого порядку (двох факторів) і одна – взаємодії другого порядку (трьох факторів).

Якщо додати ще один, четвертий фактор, то з'явиться взаємодія третього порядку, а число гіпотез, що перевіряються зросте до 15. Для п'яти факторів число гіпотез, що перевіряються складе вже 31! Загальна кількість перевірених гіпотез дорівнює  $2^P - 1$ , де  $P$  – число факторів. При збільшенні числа факторів швидко зростає не тільки кількість перевірених гіпотез, але і складність інтерпретації взаємодій. Якщо, як ми побачимо далі, інтерпретація взаємодії першого порядку (двох факторів) зазвичай не складає труднощів, то взаємодія другого порядку обов'язково вимагає побудови графіків середніх значень, а інтерпретація взаємодії третього порядку і вище навряд чи взагалі

можлива.

*Вплив коваріат.* Коваріати використовуються для виключення впливу кількісної змінної на залежну змінну. Коваріату найпростіше представити як змінну, значно корелює з залежною змінною і дозволяє зменшити дисперсію останньої. Включення в аналіз коваріати за змістом означає виключення її впливу на залежну змінну. За рахунок цього дисперсія останньої зменшується, що дозволяє зробити більш очевидним вплив факторів. В нашому дослідженні в якості коваріати буде використовуватися змінна «тест\_сер.» Ця змінна, як сумарний показник інтелектуальних здібностей, в істотній мірі корелює з успішністю («тест 2»).

Якщо ми хочемо простежити вплив факторів «стать» і «хобі» на залежну змінну «тест 2» без урахування впливу фактору «тест\_сер», то останній необхідно включити в аналіз як коваріати. Поява коваріати не впливає на описові статистики, проте може змінити суму квадратів (як правило, в меншу сторону) і величину f-критерію (як в меншу, так і, можливо, в більшу сторону).

#### *Покрокові алгоритми обчислень*

Для проведення багатофакторного дисперсійного аналізу спочатку необхідно виконати три підготовчих кроки.

**Крок 1.** Створіть новий файл даних або підготуйте існуючий.

**Крок 2.** Запустіть програму SPSS за допомогою значка на робочому столі або команди Пуск > Програми > SPSS for Windows > SPSS 11.5 for Windows (Start > Programs > SPSS for Windows > SPSS 11.5 for Windows) головного меню Windows.

У відкритому після запуску програми діалоговому вікні SPSS for Windows клацніть на кнопці Cancel (Скасувати).

Після виконання цього кроку на екрані з'явиться вікно редактора даних SPSS.

**Крок 3.** Відкрийте файл даних, з яким ви маєте намір працювати. Якщо він розташований в цій папці, то виконайте наступні дії:

1. Виберіть в меню File (Файл) команду Open > Data (Відкриття > Дані) або клацніть на кнопці Open File (Відкриття файлу) панелі інструментів.
2. У діалоговому вікні двічі клацніть по імені файлу або введіть його з клавіатури і клацніть на кнопці OK.

Незалежно від того, чи була відкрита програма SPSS тільки-но або якісь процедури вже виконувалися, у верхній частині головного вікна має бути рядок меню. Поки рядок меню наявний на екрані, доступні всі команди аналізу даних. При цьому вікно з даними бачити не обов'язково.

При роботі з таблицями результатів чи під час редагування діаграм деякі пункти меню можуть зникати або змінюватися. Щоб повернутися до головного вікна і стандартної панелі меню, клацніть на кнопці згортання або відновлення поточного вікна.

Після завершення кроку 3 на екрані має з'явитися вікно редактора даних з рядком меню.

**Крок 4.** В меню Analyze (Аналіз) виберіть команду General Linear Model > Одновимірний: Загальні лінійні моделі > Одновимірний аналіз.

В меню General Linear Model (Загальні лінійні моделі), крім команди Одновимірний (Одновимірний аналіз), наявні ще три команди: Multivariate (Багатовимірний аналіз), Repeated Measures (Повторні вимірювання) і Variance Components (Компоненти дисперсії). Команда Multivariate (Багатовимірний аналіз) використовується для багатовимірного дисперсійного і коваріаційного аналізів.

Команда Repeated Measures (Повторні вимірювання) може застосовуватися як для однофакторного, так і для багатофакторного аналізу. Команда Variance Components (Компоненти дисперсії) має дещо більшу гнучкість, порівняно з попередніми.

У лівій частині діалогового вікна Одновимірний (Одновимірний аналіз) розташовано список всіх доступних змінних файлу даних, а праворуч знаходяться 5 полів, за допомогою яких задаються основні параметри дисперсійного аналізу:

- поле Dependent Variable (Залежна змінна) призначене для вказівки єдиною залежної змінної аналізу;
- поле Fixed factor(s) (Постійні фактори) – для імен незалежних змінних, або факторів;
- поле Covariate(s) (Коваріати) – для зазначення імен коваріат.

Кнопки Contrasts (Контрасти) і Post Hoc (Постфактум) діють практично ідентично своїм аналогам з діалогового вікна One-Way ANOVA (Однофакторний дисперсійний аналіз).

Особливе значення має кнопка Plots (Діаграми), що дозволяє будувати графіки середніх значень. Для побудови графіків у програмі SPSS передбачено окреме меню Graphs (Графіки), але у зв'язку з тим, що графіки середніх значень вкрай необхідні для інтерпретації взаємодій факторів, в основні діалогового вікна команди підміню General Linear Model (Загальні лінійні моделі) включена кнопка Plots (Діаграми), що дозволяє будувати графіки середніх, не виходячи з діалогового вікна налаштування параметрів аналізу.

*Двофакторний дисперсійний аналіз.* Після завдання залежною змінною «тест 2» необхідно задати незалежні змінні, по черзі розміщуючи їх в поле Fixed factor(s) (Постійні фактори).

Коли імена всіх змінних, які беруть участь в аналізі, визначені, можна перейти до завдання параметрів виконуваних дій. Для цього клацніть на кнопці Options (Параметри). На екрані з'явиться діалогове вікно Univariate : Options (Одновимірний аналіз : Параметри). Нас буде цікавити група прапорців Display (Відображати), що дозволяє задати список включених у висновок велич. Так, прапорець Descriptive statistics (Описові статистики) дозволяє вивести для кожного елементу таблиці спряженості градацій факторів, середні значення, стандартні відхилення і розміри вибірок. Нерідко використовується прапорець Estimates of effect size (Оцінка величини ефекту), що включає в дані, які виводяться, величину ефекту, призначену для оцінки впливу кожної незалежної змінної.

Наступний приклад являє собою практичну реалізацію описаних дій.

**Крок 5.** Після виконання кроку 4 має бути відкрито діалогове

вікно Одновимірний аналіз.

1. Клацніть спочатку на змінну «тест 2», щоб виділити її, а потім – на кнопку вгорі зі стрілками, щоб перемістити змінну в поле Dependent Variable (Залежна змінна);

2. Клацніть на змінну «стать», щоб виділити її, а потім – на другу зверху кнопку зі стрілкою, щоб перемістити змінну в список Fixed Factor(s) (Постійні фактори);

3. Повторіть попередній крок для змінної «хобі»;

4. Клацніть на кнопку Options (Параметри), щоб відкрити діалогове вікно Одновимірний : Options (Одновимірний аналіз : Параметри);

5. Установіть прапорці Descriptive statistics і Estimates of effect size, а потім клацніть на кнопку Continue, щоб повернутися в діалогове вікно Одновимірний;

6. Клацніть на кнопку ОК, щоб відкрити вікно виведення;

7. В наступному покроковому алгоритмі ми реалізуємо двофакторний дисперсійний аналіз з урахуванням коваріати. Всі змінні залишаться тими ж, але додатково в якості коваріати включимо змінну «тест\_сер». Це дозволить порівняти результати дисперсійного аналізу з урахуванням і без урахування впливу коваріати.

**Крок 5а.** Після виконання кроку 4 має відкритися діалогове вікно Одновимірний аналіз. Якщо ви вже встигли попрацювати з цим вікном, клацніть на кнопку Reset (Скидання). Далі робіть наступне:

1. Клацніть спочатку на змінну «тест 2», щоб виділити її, а потім – на кнопку вгорі зі стрілками, щоб перемістити змінну в поле Dependent Variable (Залежна змінна);

2. Клацніть спочатку на змінну «стать», щоб виділити її, а потім – на другий зверху кнопки зі стрілкою, щоб перемістити змінну в список Fixed Factor(s) (Постійні фактори);

3. Повторіть попередній крок для змінної «хобі»;

4. Клацніть спочатку на змінну «тест\_сер.», щоб виділити її, а потім – на другий знизу кнопки зі стрілкою, щоб перемістити змінну в Covariates;

5. Клацніть на кнопку Options (Параметри), щоб відкрити діалогове вікно Одновимірний: Options (Одновимірний аналіз : Параметри);

6. Установіть прапорці Descriptive statistics (Описові статистики) і Estimates of effect size (Оцінка величини ефекту), а потім клацніть на кнопку Continue (Продовжити), щоб повернутися в діалогове вікно Одновимірний.

7. Клацніть на кнопку ОК, щоб відкрити вікно виведення.

Результати виконання кроку 5а будуть відрізнятися від результатів виконання кроку 5 тільки урахуванням коваріати «тест\_сер.».