**Лекція 14**

**Однофакторний дисперсійний АНАЛІЗ**

Метод дисперсійного аналізу має загальноприйняте позначення ANOVA - короткий написання терміна ANALISIS OF VARIANCE. У психологічних дослідженнях найчастіше застосовуються однофакторний ANOVA і двохфакторну ANOVA, рідше інші види дисперсійного аналізу.

За допомогою однофакторного ANOVA перевіряються статистичні гіпотези про відмінності середніх значень деякої властивості у вибірці, що класифікована за однією основою, так званому одному фактору. Тобто перевіряються статистичні гіпотези про відмінності середніх значень однієї змінної, що має три і більше градацій одного фактора.

Особливості дисперсійного аналізу ANOVA полягають в наступному:

– змінна (властивість), що порівнюється, повинна бути виміряна в метричній шкалі;

– порівнювані вибірки, що відповідають різним градаціях фактора, повинні мало відрізнятися за чисельністю, або стандартні відхилення порівнюваних вибірок відрізняються незначно;;

– відповідність розподілів частот для кожної градації фактора відповідає нормальному розподілу.

Однофакторний дисперсійний аналіз (ANOVA) можна здійснювати за допомогою Excel. Для двохфакторну дисперсійного аналізу (ANOVA) більш ефективно застосувань спеціальних статистичних комп'ютерних програм (STATISTIKA, SPSS та ін.). Тому розглянемо застосування однофакторного методу ANOVA в Excel.

Однофакторний дисперсійний аналіз (ANOVA) використовується для дослідження відмінності середніх значень однієї змінної (властивості) для трьох і більше вибірок, виділених із загальної вибірки шляхом класифікації її за однією основою (одному фактору).

За допомогою однофакторного ANOVA перевіряється гіпотеза Н0: «Середні трьох вибірок статистично значуще не відрізняються». Якщо гіпотеза Н0 відхиляється, то приймається гіпотеза Н1: «Середні принаймні двох вибірок статистично значимо відрізняються (р ≤ α)»

**Приклад вибірок, для яких можна застосувати однофакторний ANOVA**

Виміряно час читання (в секундах) одного і того ж тексту у респондентів різних вікових груп: до 22 років, від 23 до 26 років, від 27 років. Результати вимірювання представлені в протоколі.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Вік | Час |  | № | Вік | Час |  | № | Вік | Час |
| 1 | 20 | 35 |  | 1 | 23 | 41 |  | 1 | 27 | 46 |
| 2 | 20 | 37 |  | 2 | 23 | 40 |  | 2 | 27 | 40 |
| 3 | 21 | 43 |  | 3 | 24 | 42 |  | 3 | 28 | 40 |
| 4 | 21 | 35 |  | 4 | 24 | 39 |  | 4 | 28 | 39 |
| 5 | 21 | 35 |  | 5 | 24 | 38 |  | 5 | 28 | 38 |
| 6 | 21 | 39 |  | 6 | 25 | 37 |  | 6 | 29 | 37 |
| 7 | 21 | 40 |  | 7 | 25 | 37 |  | 7 | 29 | 60 |
| 8 | 22 | 34 |  | 8 | 25 | 38 |  | 8 | 29 | 39 |
| 9 | 22 | 34 |  | 9 | 25 | 40 |  | 9 | 30 | 40 |
| 10 | 22 | 34 |  | 10 | 25 | 41 |  | 10 | 31 | 38 |
| 11 | 22 | 35 |  | 11 | 25 | 39 |  |   |  |   |
| 12 | 22 | 36 |  |   |   |   |  |   |   |   |
| 13 | 22 | 37 |  |   |   |   |  |   |   |   |

***Питання, відповідь на який можна знайти за допомогою однофакторного аналізу***

Чи є статистично значущі відмінності часу, витраченого на читання тексту респондентами різних вікових груп: до 22 років, від 23 до 26 років, від 27 років?

***Алгоритм однофакторного ANOVA (Excel)***

1. На окремому аркуші в Excel створюється таблиця, стовпцями якої присвоюються назви відповідні градаціях фактора. З протоколу в створену таблицю переносяться відповідні значення змінної (виміряної властивості).

2. Знайдіть в групі команд Дані команду Аналіз даних.

***Якщо немає, тоді::***

1. *На вкладке****Файл****выберите команду****Параметры****, а затем — категорию****Надстройки****.*

*Если вы используете Excel 2007, нажмите кнопку****Microsoft Office**** и выберите Параметры****Excel.***

1. *В раскрывающемся списке****Управление****выберите пункт****Надстройки Excel****и нажмите кнопку****Перейти****.*

*Если вы используете Excel для Mac, в строке меню откройте вкладку****Средства****и в раскрывающемся списке выберите пункт****Надстройки для Excel****.*

1. *В диалоговом окне****Надстройки****установите флажок****Пакет анализа****, а затем нажмите кнопку****ОК****.*
	* *Если****Пакет анализа****отсутствует в списке поля****Доступные надстройки****, нажмите кнопку****Обзор****, чтобы выполнить поиск.*
	* *Если выводится сообщение о том, что пакет анализа не установлен на компьютере, нажмите кнопку****Да****, чтобы установить его.*



3. В інструментах аналізу знайдіть Однофакторний дисперсійний аналіз. Натисніть ОК.



4. У вікно Вхідний інтервал внесіть коди створеної таблиці.

5. У вікні Мітки в першому рядку поставте знак «v». Натисніть ОК.

6. У новому вікні з'явиться таблиця результатів однофакторного ANOVA.

7. Знайдіть у таблиці стовпець P-Значення, в якому відзначено шуканий рівень значущості відмінностей.

***Приклад однофакторного ANOVA для зв'язаних вибірок (Excel)***

У протоколі наведені результати вимірювання часу (в секундах) читання одного і того ж тексту респондентами.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Вік | Час |  | № | Вік | Час |  | № | Вік | Час |
| 1 | 20 | 35 |  | 1 | 23 | 41 |  | 1 | 27 | 46 |
| 2 | 20 | 37 |  | 2 | 23 | 40 |  | 2 | 27 | 40 |
| 3 | 21 | 43 |  | 3 | 24 | 42 |  | 3 | 28 | 40 |
| 4 | 21 | 35 |  | 4 | 24 | 39 |  | 4 | 28 | 39 |
| 5 | 21 | 35 |  | 5 | 24 | 38 |  | 5 | 28 | 38 |
| 6 | 21 | 39 |  | 6 | 25 | 37 |  | 6 | 29 | 37 |
| 7 | 21 | 40 |  | 7 | 25 | 37 |  | 7 | 29 | 60 |
| 8 | 22 | 34 |  | 8 | 25 | 38 |  | 8 | 29 | 39 |
| 9 | 22 | 34 |  | 9 | 25 | 40 |  | 9 | 30 | 40 |
| 10 | 22 | 34 |  | 10 | 25 | 41 |  | 10 | 31 | 38 |
| 11 | 22 | 35 |  | 11 | 25 | 39 |  |   |  |   |
| 12 | 22 | 36 |  |   |   |   |  |   |   |   |
| 13 | 22 | 37 |  |   |   |   |  |   |   |   |

Потрібно встановити рівень статистичної значущості відмінностей середніх значень часу читання текстів респондентами різних вікових груп: до 22 років, від 23 до 26 років, від 27 років.

Фактором є вік респондентів.

Фактор має три градації: до 22 років, від 23 до 26 років, від 27 років.

Досліджуваної змінної (властивістю) є час, витрачений респондентом на читання тексту.

Замінимо в таблиці стовпець «Вік» на стовпець «Фактор».

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | фактор | Час |  | № | фактор | Час |  | № | фактор | Час |
| 1 | до 22 | 35 |  | 1 | від 23 до 26 | 41 |  | 1 | більше 27 | 46 |
| 2 | до 22 | 37 |  | 2 | від 23 до 26 | 40 |  | 2 | більше 27 | 40 |
| 3 | до 22 | 43 |  | 3 | від 23 до 26 | 42 |  | 3 | більше 27 | 40 |
| 4 | до 22 | 35 |  | 4 | від 23 до 26 | 39 |  | 4 | більше 27 | 39 |
| 5 | до 22 | 35 |  | 5 | від 23 до 26 | 38 |  | 5 | більше 27 | 38 |
| 6 | до 22 | 39 |  | 6 | від 23 до 26 | 37 |  | 6 | більше 27 | 37 |
| 7 | до 22 | 40 |  | 7 | від 23 до 26 | 37 |  | 7 | більше 27 | 60 |
| 8 | до 22 | 34 |  | 8 | від 23 до 26 | 38 |  | 8 | більше 27 | 39 |
| 9 | до 22 | 34 |  | 9 | від 23 до 26 | 40 |  | 9 | більше 27 | 40 |
| 10 | до 22 | 34 |  | 10 | від 23 до 26 | 41 |  | 10 | більше 27 | 38 |
| 11 | до 22 | 35 |  | 11 | від 23 до 26 | 39 |  |  |  |  |
| 12 | до 22 | 36 |  |  |  |  |  |  |  |  |
| 13 | до 22 | 37 |  |  |  |  |  |  |  |  |

На окремому лист в Excel створимо таблицю зі стовпцями: до 22 років, від 23 до 26 років, від 27 років (назви градацій фактора). З протоколу в створену таблицю перенесемо відповідні значення змінної (час, витрачений на читання тексту).

|  |  |  |
| --- | --- | --- |
| **до 22** | **від 23 до 26** | **більше 27** |
| 35 | 41 | 46 |
| 37 | 40 | 40 |
| 43 | 42 | 40 |
| 35 | 39 | 39 |
| 35 | 38 | 38 |
| 39 | 37 | 37 |
| 40 | 37 | 60 |
| 34 | 38 | 39 |
| 34 | 40 | 40 |
| 34 | 41 | 38 |
| 35 | 39 |  |
| 36 |  |  |
| 37 |  |  |

Знайдемо в групі команд Дані команду Аналіз даних. Натиснемо ОК.

В Інструментах аналізу знайдемо Однофакторний дисперсійний аналіз. Натиснемо ОК.

У вікно Вхідний інтервал внесемо коди створеної таблиці шляхом виділення цієї таблиці. У вікні Мітки в першому рядку поставимо знак «v». Натиснемо ОК.

У новому вікні з'явиться таблиця результатів однофакторного ANOVA.

Знайдемо в таблиці стовпець P-Значення, в якому відзначено шуканий рівень значущості відмінностей: α = 0,02 (число округлено до двох знаків після коми).

**Статистичний висновок.**

Оскільки α = 0,02, т.е. р ≤ 0,02, приймається гіпотеза Н1: Середні принаймні двох вибірок статистично значимо відрізняються (р ≤ 0,02).

**Змістовний висновок.**

Виявлено статистично значущі (р ≤ 0,02) відмінності середніх значень часу, витраченого на читання текстів, у респондентів різних вікових груп.

