|  |  |  |  |
| --- | --- | --- | --- |
| МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені Тараса Шевченка  ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  **Кафедра програмних систем і технологій**        Дисципліна  **«Емпіричні методи програмної інженерії»**      **Практична робота № 5**  **“**ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ**”** | | | |
| **Виконав:** | Гоша Д. О | **Перевірила**: | Юрчук Ірина Аркадіївна |
| Група | ІПЗ-23 | Дата перевірки |  |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |

Мета: Навчитися формулювати гіпотези стосовно даних, які отримані при вимірювання, та освоїти статистичні методи для їх прийняття чи спростування.

# Завдання

1. Сформулювати гіпотези відносно ідентичності 3 метрик у двох предметних областях (по 10 проектів) та 5 проектів у кожній предметній області окремо.
2. Застосувати статистичні тести з метою прийняття чи відхилення сформульованих гіпотез.
3. Надати висновки з отриманих результатів вказавши помилку І роду (𝛼).

# Варіант 4

# Виконання лабораторної роботи

Першою гіпотезою, яка буде сформульована, є **гіпотеза відносно ідентичності метрик** NOC, NOP, NOM відносно різних груп (дрібних та великих проектів), адже ці метрики за теорією є більш незалежні від мови програмування, аніж інші: кількість рядків коду хоча й може бути менше для проектів другої шрупи буде значно більше можливо навіть у десятки разів від малих проектів. Адже Великий проект має багато додаткового функціоналу. Але це ніяк не вплине на залежність рядків коду, наприклад розділену на кількість методів, отже якщо порівнювати відношення кількості рядків коду на кількість методів отримаємо середній розмір методу, і за стандартами програмування метод має бути не завеликим і не замалим в будь якому проекті.

Наступною метрикою буде кількість пакетів, хоча й може показувати кращу розмежованість за логікою відповідно до архітектурних стиль, які краще впроваджені у масштабних проектах, але також кількість пакетів може свідчити про надмірність цього розподілу, що порушує архітектуру, чим іноді відрізняються маленькі проекти.

Кількість методів класів незалежно від розміру проекту має буди статично у будь якому проекті. Звичайною менших проектів спресовують методи менші за логікою та функціоналом у один клас. Більш вагомою буде сама архітектури та підхід до розробки ПЗ , уніфіковані стандарти самої студії розробки.

Отже, як можемо побачити, ці метрики можуть бути більше як з питань покращенню коду, так і з питань його погіршення, а отже вони можуть бути ідентичні для обох груп.

Для підтвердження чи спростування цієї гіпотези будемо використовувати U-критерій Манна-Уітні, який полягає в складанні єдиного варіаційного ряду з обох вибірок та визначення рангу для кожного окремо елементу цього ряду з подальшим визначенням суми рангів окремо для кожної з груп. Після цього визначається більша з цих сум (𝑇х) та кількість елементів для групи, що відповідає цій сумі (𝑛𝑥). Таким чином, застосовується така формула для визначення значення U-критерію:

Після цього для підтвердження гіпотези цей критерій перевіряється відносно критичного значення 𝑈𝑐𝑟𝑖𝑡 = 23 (за таблицею для 𝑛𝑚𝑖𝑛 = 10, 𝑛𝑚𝑎𝑥 = 10) на те, що він більший за нього. Важливо зазначити, що обрання критичного значення здійснювалося з рівнем значущості 𝛼 = 0.05, адже існує певна кількість факторів, які викликають аномальні результати для метрик, тому потрібен дещо більший діапазон допустимих значень для прийняття гіпотези. Для здійснення таких обчислень в середовищі Excel потрібно виконати такі операції:

1. Визначення рангів метрик: використання функції RANK.AVG для варіаційного ряду.
2. Визначення суми рангів: використання функції SUM для першої та другої суми відповідно.
3. Знаходження U-критерію за формулою.
4. Перевірка гіпотези: порівняння U-критерію з 𝑈𝑐𝑟𝑖𝑡

3 Таким чином, застосуємо цей алгоритм для дослідження обраних метрик на ідентичність:

1. **NOC**
   1. Таблиця "значення-ранг".

|  |  |  |  |
| --- | --- | --- | --- |
| Маленькі проекти | Ранг | Масштабні проекти | Ранг |
| 199 | 20 | 62274 | 4 |
| 43921 | 5 | 2475 | 14 |
| 20098 | 9 | 983 | 16 |
| 21220 | 8 | 3160 | 13 |
| 611 | 17 | 34563 | 6 |
| 5541 | 11 | 67345 | 3 |
| 5043 | 12 | 23453 | 7 |
| 7064 | 10 | 341 | 18 |
| 1408 | 15 | 67700 | 2 |
| 67923 | 1 | 234 | 19 |
| **Сума** | **108** | **Сума** | **102** |

* 1. Аналіз гіпотези: для цього варіаційного ряду був отриманий Uкритерій  **47** (, який є більшим за критичне значення 𝑈𝑐𝑟𝑖𝑡 = 23 за рівню значущості 𝛼 = 0.05, а отже можемо підтвердити гіпотезу про ідентичність цієї метрики відносно розглянутих груп.

1. **NOM**
   1. Таблиця "значення-ранг".

|  |  |  |  |
| --- | --- | --- | --- |
| Маленькі проекти | Ранг | Масштабні проекти | Ранг |
| 3,2 | 18 | 4,3 | 14 |
| 9,2 | 1,5 | 6,8 | 10 |
| 5,7 | 12 | 9,2 | 1,5 |
| 9,1 | 3 | 5,9 | 11 |
| 8,3 | 4 | 5,2 | 13 |
| 7 | 8 | 8 | 5 |
| 4,2 | 15 | 7,8 | 6 |
| 3,6 | 16,5 | 2,9 | 19 |
| 7,1 | 7 | 3,6 | 16,5 |
| 6,9 | 9 | 2,6 | 20 |
| **Сума** | **94** | **Сума** | **116** |

* 1. Аналіз гіпотези: для цього варіаційного ряду був отриманий Uкритерій  **39** (, який є більшим за критичне значення 𝑈𝑐𝑟𝑖𝑡 = 23 за рівню значущості 𝛼 = 0.05, а отже можемо підтвердити гіпотезу про ідентичність цієї метрики відносно розглянутих груп.

1. **NOP**
   1. Таблиця "значення-ранг".

|  |  |  |  |
| --- | --- | --- | --- |
| Маленькі проекти | Ранг | Масштабні проекти | Ранг |
| 1,62 | 20 | 6,07 | 8 |
| 4,86 | 11 | 9,05 | 4 |
| 3,84 | 16 | 19 | 1 |
| 6,03 | 9 | 4,69 | 12 |
| 2,46 | 19 | 7 | 6 |
| 7,88 | 5 | 4,54 | 13 |
| 2,95 | 17 | 4,42 | 14 |
| 10,24 | 3 | 6,51 | 7 |
| 2,5 | 18 | 4,35 | 15 |
| 10,5 | 2 | 5,12 | 10 |
| **Сума** | **120** | **Сума** | **90** |

* 1. Аналіз гіпотези: для цього варіаційного ряду був отриманий Uкритерій  **35** (, який є більшим за критичне значення 𝑈𝑐𝑟𝑖𝑡 = 23 за рівню значущості 𝛼 = 0.05, а отже можемо підтвердити гіпотезу про ідентичність цієї метрики відносно розглянутих груп.

**Гіпотеза відносно ідентичності проектів.**

Другою гіпотезою, яка буде сформульована, є гіпотеза відносно ідентичності 5 проектів однієї групи, адже саме їхня ідентичність може стверджувати або спростовувати певні припущення, які були висунуті під час першої практичної роботи, про функціональність, внутрішній устрій цих проектів. Таким чином, були обрані такі проекти:

1. Дочірні проекти:
   1. Blazored TextEditor (https://github.com/Blazored/TextEditor)
   2. AvalonEdit (https://github.com/icsharpcode/AvalonEdit)
   3. Limetext (https://github.com/limetext/lime/tree/old\_master)
   4. FastColoredTextBox (https://github.com/PavelTorgashov/FastColoredTextBox)
   5. АvaloniaEdit (https://github.com/AvaloniaUI/AvaloniaEdit)
2. Масштабні проекти:
   1. Vi / Vim
   2. Gedit
   3. Nano
   4. GNU
   5. Kate/Kwrite

Для підтвердження чи спростування цієї гіпотези будемо використовувати критерій Краскала-Уолліса, який полягає в складанні варіаційного ряду з усіх метрик усіх проектів, які досліджуються, та визначення рангу для кожного окремого елементу ряду з подальшим визначенням суми рангів (𝑇𝑖 ) окремо для кожної групи (проекту). Таким чином, застосовується така формула для отримання значення критерію:

Після цього для підтвердження гіпотези цей критерій перевіряється відносно критичного значення 𝜒 2 = 9.49 (за таблицею для 𝑑𝑓 = 𝑘 − 1 = 5 − 1 = 4) на те, що він менший за нього.

Важливо зазначити, що обрання критичного значення здійснювалося з рівнем значущості 𝛼 = 0.05, адже існує певна к-сть факторів, які викликають аномальні результати для метрик, тому потрібен дещо більший діапазон допустим значень для прийняття гіпотези. Для здійснення таких обчислень в середовищі Excel потрібно виконати такі операції:

* Визначення рангів елементів ряду: використання функції RANK.AVG для варіаційного ряду.
* Визначення суми рангів: використання функції SUM для кожної групи елементів (проекту) відповідно.
* Знаходження H-критерію за формулою.
* Перевірка гіпотези: порівняння H-критерію з 𝜒.

Таким чином, застосуємо цей алгоритм для дослідження обраних проектів на ідентичність:

1. **Дрібні проекти**
   1. Таблиця "значення-ранг-сума"

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | | 2 | | 3 | | 4 | | 5 | |
| ЗН. | РН. | ЗН. | РН. | ЗН. | РН. | ЗН. | РН. | ЗН. | РН. |
| NOC | 199 | 5 | 43921 | 1 | 20098 | 3 | 21220 | 2 | 611 | 4 |
| NOM | 3,2 | 27 | 9,2 | 10 | 5,7 | 20 | 9,1 | 11 | 8,3 | 16 |
| NDD | 3 | 29 | 9 | 13,5 | 7 | 17,5 | 9 | 13,5 | 7 | 17,5 |
| NOP | 1,62 | 40 | 4,86 | 22 | 3,84 | 24 | 6,03 | 19 | 2,46 | 32 |
| CALL | 1,75 | 39 | 2,43 | 33 | 2,15 | 36 | 2,72 | 31 | 2,32 | 34 |
| WMC | 3 | 29 | 118 | 6 | 96 | 7 | 66 | 9 | 68 | 8 |
| TCC | 2,16 | 35 | 3,24 | 26 | 3,4 | 25 | 2,13 | 37 | 4,17 | 23 |
| PNAS | 2 | 38 | 9 | 13,5 | 9 | 13,5 | 5 | 21 | 3 | 29 |
| Сума |  | **242** |  | **125** |  | **146** |  | **143,5** |  | **163,5** |

* 1. Аналіз гіпотези: для цього варіаційного ряду був отриманий Hкритерій **7,63673**, який є приблиздним до критичного значення 𝜒 2 = 7.78 за рівню значущості 𝛼 = 0.1, а отже можемо стверджувати, що обрані проекти є нвближено ідентичні з точки зору емпіричних характеристик, що підтверджує певні припущення, які були висунуті під час практичної роботи.

1. **Масштабні проекти**
   1. Таблиця "значення-ранг-сума"

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | | 2 | | 3 | | 4 | | 5 | |
| ЗН. | РН. | ЗН. | РН. | ЗН. | РН. | ЗН. | РН. | ЗН. | РН. |
| NOC | 234 | 5 | 67700 | 1 | 341 | 4 | 23453 | 3 | 67345 | 2 |
| NOM | 2,6 | 30 | 3,6 | 27 | 2,9 | 29 | 7,8 | 15 | 8 | 13 |
| NDD | 7 | 18 | 5 | 23 | 1 | 40 | 7 | 18 | 8 | 13 |
| NOP | 5,12 | 22 | 4,35 | 26 | 6,51 | 21 | 4,42 | 25 | 4,54 | 24 |
| CALL | 2,05 | 34 | 1,98 | 35 | 1,51 | 38 | 1,94 | 36 | 2,21 | 32 |
| WMC | 51 | 8 | 43 | 9,5 | 118 | 6 | 89 | 7 | 43 | 9,5 |
| TCC | 1,45 | 39 | 2,28 | 31 | 2,14 | 33 | 3,1 | 28 | 1,7 | 37 |
| PNAS | 7 | 18 | 9 | 11 | 7 | 18 | 7 | 18 | 8 | 13 |
| Сума |  | **174** |  | **163,5** |  | **189** |  | **150** |  | **143,5** |

* 1. Аналіз гіпотези: для цього варіаційного ряду був отриманий Hкритерій **1,226982** який є значно меньшим за критичного значення 𝜒 2 = 9.84 за рівню значущості 𝛼 = 0.05, а отже можемо стверджувати, що обрані проекти є майже ідентичні з точки зору емпіричних характеристик, що підтверджує певні припущення, які були висунуті під час практичної роботи.

**Висновки**

Під час цієї лабораторної роботи було проведено дослідження на перевірку на ідентичність метрик та проектів. Для цього була висунута гіпотеза щодо ідентичності 3 метрик (NOC, NOP, NOM) відносно різних груп (великих та малих), що викликано певною незалежністю цих метрик від мови програмування, а також була висунута гіпотеза про ідентичність 5 проектів для кожної з групи, що викликано перевіркою припущень, які були висунуті раніше. По-перше, при перевірці гіпотез щодо ідентичності метрик були отримані такі результати:

1. NOC– підтвердження гіпотези. Справді, кількість рядків коду не так сильно відрізняється, що можна було б стверджувати, що маємо якусь достатню різницю між групами, але важливо зазначити, що певна різниця все одно присутня, що було розглянуто в практичні роботі 7.

2. NOP – підтвердження гіпотези. Справді, кількість пакетів також не відрізняються суттєво, але це здебільшого викликано не правильною структуризацією для обох груп, адже спостерігається певна алогічність та надлишковість у пакетуванні Java-проектів.

3. NOM – спростування гіпотези. Справді, як було доведено в практичній роботі 7, маємо суттєву різницю між проектами обох груп за цією метрикою, що викликано різним підходом до розробки архітектуру та стандартами і класифікаціями, що вказують яка кількість методів додається у класу, в залежності від компанії розробника ПЗ.

По-друге, при перевірці гіпотез щодо ідентичності проектів були підтверджені обидві гіпотези як щодо великих проектів , так і щодо малих проектів . Цей результат дозволяє підтвердити припущення щодо функціональності, внутрішнього устрою та структури обраних проектів, що були висунуті раніше, зокрема під час лабораторної 1. Отже, після виконання цієї лабораторної роботи була отримана інформація щодо ідентичності певних метрик та проектів, що дозволило підтвердити попередні дослідження та деталізувати їх з інших аспектів..