|  |  |  |
| --- | --- | --- |
| Министерство науки и высшего образования  Российской Федерации | | |
| Федеральное государственное бюджетное  образовательное учреждение высшего образования | | |
| «Новосибирский государственный технический университет» | | |
|  | | |
| Кафедра теоретической и прикладной информатики | | |
| Занятие №7 | | |
| по дисциплине «Методы принятия оптимальных решений» | | |
| **Применение критериев проверки гипотез об однородности (законов, средних значений, дисперсий).** | | |
|  | | |
|  | Факультет: | ПМИ |
| Группа: | ПМИ-12 |
|  |  |
| Студенты: | Попов Семён |
|  | Кораблёв Кирилл  Швадченко Артём |
|  |  |
| Преподаватель: | Лемешко Борис Юрьевич |
|  |  |
|  |  |  |
| Новосибирск  2024 | | |

Задание 1:

Смоделировать 2 выборки в соответствии с нормальным законом объемом n = 1000 одна с параметрами сдвига и масштаба (0, 1), вторая – с параметрами (0.1, 1), то есть математическое ожидание второй отличается на 10% от стандартного отклонения.

Последовательно, меняя объем выборки n = 25, 50, 100, 200, 500, 1000 (в тексте анализируемых выборок менять объём и перегружать выборки), проследите, как меняется достигаемый уровень значимости при проверке гипотезы об однородности средних по следующим критериям: Стьюдента, сравнения средних при неизвестных и неравных дисперсиях, Краскела-Уаллиса, Манна-Уитни-Уилкоксона.

Зафиксируйте результаты проверок в таблице.

Оцените, примерно какой объем выборок потребуется, чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода α = 0.1, 0.05, 0.01.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Критерий | Гипотеза | Статистика | p-value |
| 25 | При неизвестных и неравных дисперсиях | НЕ ОТКЛОНЯЕТСЯ | -0.885446 | 0.388 |
| Критерий Стьюдента | НЕ ОТКЛОНЯЕТСЯ | -0.885446 | 0.388 |
| Краскела-Уаллиса | НЕ ОТКЛОНЯЕТСЯ | 0.712376 | 0.399 |
| Манна-Уитни-Уилкоксона | НЕ ОТКЛОНЯЕТСЯ | 0.844024 | 0.412 |
|  | | | | |
| 50 | При неизвестных и неравных дисперсиях | НЕ ОТКЛОНЯЕТСЯ | -0.0884407 | 0.93 |
| Критерий Стьюдента | НЕ ОТКЛОНЯЕТСЯ | -0.0884407 | 0.93 |
| Краскела-Уаллиса | НЕ ОТКЛОНЯЕТСЯ | 0.000760396 | 0.982 |
| Манна-Уитни-Уилкоксона | НЕ ОТКЛОНЯЕТСЯ | -0.0275753 | 0.98 |
|  | | | | |
| 100 | При неизвестных и неравных дисперсиях | НЕ ОТКЛОНЯЕТСЯ | -0.623623 | 0.556 |
| Критерий Стьюдента | НЕ ОТКЛОНЯЕТСЯ | -0.623623 | 0.556 |
| Краскела-Уаллиса | НЕ ОТКЛОНЯЕТСЯ | 0.150931 | 0.733 |
| Манна-Уитни-Уилкоксона | НЕ ОТКЛОНЯЕТСЯ | 0.388499 | 0.706 |
|  | | | | |
| 200 | При неизвестных и неравных дисперсиях | НЕ ОТКЛОНЯЕТСЯ | -1.2986 | 0.214 |
| Критерий Стьюдента | НЕ ОТКЛОНЯЕТСЯ | -1.2986 | 0.214 |
| Краскела-Уаллиса | НЕ ОТКЛОНЯЕТСЯ | 1.3554 | 0.246 |
| Манна-Уитни-Уилкоксона | НЕ ОТКЛОНЯЕТСЯ | 1.16422 | 0.264 |
|  | | | | |
| 500 | При неизвестных и неравных дисперсиях | ОТКЛОНЯЕТСЯ | -1.74023 | 0.096 |
| Критерий Стьюдента | ОТКЛОНЯЕТСЯ | -1.74023 | 0.096 |
| Краскела-Уаллиса | ОТКЛОНЯЕТСЯ | 4.10822 | 0.045 |
| Манна-Уитни-Уилкоксона | ОТКЛОНЯЕТСЯ | 2.02687 | 0.056 |
|  | | | | |
| 1000 | При неизвестных и неравных дисперсиях | ОТКЛОНЯЕТСЯ | -2.25038 | 0.024 |
| Критерий Стьюдента | ОТКЛОНЯЕТСЯ | -2.25038 | 0.024 |
| Краскела-Уаллиса | ОТКЛОНЯЕТСЯ | 5.39155 | 0.022 |
| Манна-Уитни-Уилкоксона | ОТКЛОНЯЕТСЯ | 2.32197 | 0.02 |

Гипотеза об однородности средних по всем критериям отклоняется при n = 500 и n = 1000.

|  |  |  |  |
| --- | --- | --- | --- |
| α | 0.1 | 0.05 | 0.01 |
| n | 500 | 650 | >1000 |

Задание 2:

Смоделировать 2 выборки в соответствии с нормальным законом объемом n = 1000 одна с параметрами (0, 1), вторая с параметрами (0, 1.1), то есть стандартное отклонение второй отличается на 10% больше чем у первой.

Последовательно, меняя объем выборки n = 25, 50, 100, 200, 500, 1000 (в тексте анализируемых выборок менять объём и перегружать выборки), проследите, как меняется достигаемый уровень значимости при проверке гипотезы об однородности дисперсий (характеристик рассеяния) по следующим критериям: параметрическим – Бартлетта, Фишера, непараметрическим – Ансари-Бредли (нормированному), Муда (нормированному), Сижеля-Тьюки (нормированному).

Зафиксируйте результаты проверок в таблице.

Оцените, примерно какой объем выборок потребуется, чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода α = 0.1, 0.05, 0.01.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Критерий | Гипотеза | Статистика | p-value |
| 25 | Бартлетта | НЕ ОТКЛОНЯЕТСЯ | 0,0756376 | 0,793 |
| Фишера | НЕ ОТКЛОНЯЕТСЯ | 0,89273 | 0,78 |
| Ансари-Бредли | НЕ ОТКЛОНЯЕТСЯ | -1,00955 | 0,306 |
| Муда | НЕ ОТКЛОНЯЕТСЯ | 0,939337 | 0,344 |
| Сижела-Тьюки | НЕ ОТКЛОНЯЕТСЯ | -0,979844 | 0,336 |
|  | | | | |
| 50 | Бартлетта | НЕ ОТКЛОНЯЕТСЯ | 0,0994491 | 0,736 |
| Фишера | НЕ ОТКЛОНЯЕТСЯ | 0,913405 | 0,724 |
| Ансари-Бредли | НЕ ОТКЛОНЯЕТСЯ | 0,0275794 | 0,962 |
| Муда | НЕ ОТКЛОНЯЕТСЯ | -0,277064 | 0,774 |
| Сижела-Тьюки | НЕ ОТКЛОНЯЕТСЯ | 0,0344691 | 0,964 |
|  | | | | |
| 100 | Бартлетта | НЕ ОТКЛОНЯЕТСЯ | 0,242293 | 0,621 |
| Фишера | НЕ ОТКЛОНЯЕТСЯ | 1,1043 | 0,636 |
| Ансари-Бредли | НЕ ОТКЛОНЯЕТСЯ | 0,156383 | 0,84 |
| Муда | НЕ ОТКЛОНЯЕТСЯ | 0,142736 | 0,936 |
| Сижела-Тьюки | НЕ ОТКЛОНЯЕТСЯ | 0,14416 | 0,842 |
|  | | | | |
| 200 | Бартлетта | НЕ ОТКЛОНЯЕТСЯ | 5.88117 | 0.013 |
| Фишера | НЕ ОТКЛОНЯЕТСЯ | 1.41214 | 0.012 |
| Ансари-Бредли | НЕ ОТКЛОНЯЕТСЯ | -2.19006 | 0.03 |
| Муда | НЕ ОТКЛОНЯЕТСЯ | 2.38653 | 0.02 |
| Сижела-Тьюки | НЕ ОТКЛОНЯЕТСЯ | -2.1961 | 0.03 |
|  | | | | |
| 500 | Бартлетта | ОТКЛОНЯЕТСЯ | 6.99024 | 0 |
| Фишера | ОТКЛОНЯЕТСЯ | 1.26758 | 0 |
| Ансари-Бредли | ОТКЛОНЯЕТСЯ | -2.29622 | 0.012 |
| Муда | ОТКЛОНЯЕТСЯ | 2.28372 | 0.012 |
| Сижела-Тьюки | ОТКЛОНЯЕТСЯ | -2.29797 | 0.012 |
|  | | | | |
| 1000 | Бартлетта | ОТКЛОНЯЕТСЯ | 5.88117 | 0.001 |
| Фишера | ОТКЛОНЯЕТСЯ | 1.41214 | 0 |
| Ансари-Бредли | ОТКЛОНЯЕТСЯ | -2.19006 | 0.036 |
| Муда | ОТКЛОНЯЕТСЯ | 2.38653 | 0.018 |
| Сижела-Тьюки | ОТКЛОНЯЕТСЯ | -2.1961 | 0.036 |

Гипотеза об однородности дисперсий (статистик рассеяния) отклоняется по всем критериям при n = 500, 1000, потому что при вероятности ошибки первого рода α = 0.1 p-value оказалось меньше заданного уровн значимости.

|  |  |  |  |
| --- | --- | --- | --- |
| α | 0.1 | 0.05 | 0.01 |
| n | 500 | 700 | >1000 |

Задание 3:

Смоделировать 2 выборки, одну – в соответствии со стандартным нормальным законом объемом n = 10000, другую – по логистическому закону с параметрами сдвига и масштаба (0, 0.551328895).

Последовательно, меняя объемы выборок n = 50, 100, 500, 1000, 5000, 10000 (в тексте анализируемых выборок менять объём и перегружать выборки), проследите, как меняется достигаемый уровень значимости при проверке гипотезы об однородности законов по критериям Смирнова и Лемана-Розенблатта.

Зафиксируйте результаты проверок в таблице.

Оцените, какой примерно объем выборок потребуется, чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода α = 0.1, 0.05, 0.01.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Критерий | Гипотеза | Статистика | p-value |
| 50 | Лемана-Розенблатта | НЕ ОТКЛОНЯЕТСЯ | 0.1454 | 0.393 |
| Смирнова | НЕ ОТКЛОНЯЕТСЯ | 0.9 | 0.309 |
|  | | | | |
| 100 | Лемана-Розенблатта | НЕ ОТКЛОНЯЕТСЯ | 0.1454 | 0.422 |
| Смирнова | НЕ ОТКЛОНЯЕТСЯ | 0.9 | 0.379 |
|  | | | | |
| 500 | Лемана-Розенблатта | НЕ ОТКЛОНЯЕТСЯ | 0.09448 | 0.601 |
| Смирнова | НЕ ОТКЛОНЯЕТСЯ | 0.885438 | 0.382 |
|  | | | | |
| 1000 | Лемана-Розенблатта | НЕ ОТКЛОНЯЕТСЯ | 0.227073 | 0.243 |
| Смирнова | НЕ ОТКЛОНЯЕТСЯ | 1.02859 | 0.234 |
|  | | | | |
| 5000 | Лемана-Розенблатта | ОТКЛОНЯЕТСЯ | 1.19453 | 0 |
| Смирнова | ОТКЛОНЯЕТСЯ | 2.24 | 0 |
|  | | | | |
| 10000 | Лемана-Розенблатта | ОТКЛОНЯЕТСЯ | 1.47939 | 0 |
| Смирнова | ОТКЛОНЯЕТСЯ | 2.14253 | 0 |

Гипотеза об однородности законов отклоняется по обоим критериям при n = 5000, 10000, потому что при вероятности ошибки первого рода α = 0.1 p-value оказалось меньше заданного уровня значимости.

|  |  |  |  |
| --- | --- | --- | --- |
| α | 0.1 | 0.05 | 0.01 |
| n | 1500 | 1700 | 4500 |

Задание 4:

Кратко сформулируйте для себя выводы, вытекающие из Ваших результатов.

1. Примерные объемы выборок при проверки однородности критериев (500, 650, >1000) при соответствующих уровнях значимости (0.1, 0.05, 0.01) различны.
2. Примерные объемы выборок при проверке однородности дисперсий так же различны, однако имеют меньшую разницу, (500, 700, >1000) при соответствующих уровнях значимости (0.1, 0.05, 0.01).

При проверке об однородности законов наблюдаются объемы выборок гораздо больше, чем в предыдущих пунктах (1500, 1700, 4500) при соответствующих уровнях значимости (0.1, 0.05, 0.01).