Министерство Образования и Науки Российской Федерации

Новосибирский Государственный Технический Университет

Кафедра теоретической и прикладной информатики

ЛАБОРАТОРНАЯ РАБОТА № 7

по предмету: История и методология прикладной математики и информатики

на тему «Применение критериев проверки гипотез об однородности (законов, средних значений, дисперсий)»

#### Факультет: ПМИ

Группа: ПММ-61  
Студенты: Архипенко Е.П.

Захаров С.Б.

Кочан М.В.

Преподаватели: Лемешко Б.Ю.

Новосибирск  
 2016

**Цель занятия** – оценить возможности некоторых критериев проверки статистических гипотез по различению близких альтернатив (критериев однородности средних, однородности дисперсий, однородности законов). Проследить, как в зависимости от объемов выборок меняются результаты проверки (p-value). По результатам экспериментов оценить, примерно какой объем выборок потребуется, чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу.

1. *Смоделировать 2 выборки в соответствии с нормальным законом объемом =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  (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 .*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| n | Критерий Стьюдента | Кр. сравнения средних при неизв-х. и неравных дисперсиях | Критерий Краскела-Уаллиса | Критерий Манна-Уитни-Уилкоксона |
| 25 | p-value=0.180525 | p-value=0.180618 | p-value=0.242 | p-value=0.217919 |
| 50 | p-value=0.591123 | p-value=0.591123 | p-value=0.571874 | p-value=0.571874 |
| 100 | p-value=0.640501 | p-value=0.640501 | p-value=0.69584 | p-value=0.69584 |
| 200 | p-value=0.282896 | p-value=0.282919 | p-value=0.231949 | p-value=0.231949 |
| 500 | p-value=0.122914 | p-value=0.122918 | p-value=0.078048 | p-value=0.078048 |
| 1000 | p-value=0.32456 | p-value=0.32456 | p-value=0.189157 | p-value=0.189157 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется *n* примерно от 360 до 670 (для критериев Краскела-Уаллиса и Манна-Уитни-Уилкоксона).

1. *Смоделировать 2 выборки в соответствии с нормальным законом объемом =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  (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 .*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| n | Критерий Бартлетта | Критерий Фишера | Критерий Ансари-Бредли (норм.) | Критерий Муда (норм.) | Критерий Сижеля-Тьюки (норм.) |
| 25 | p-value=  0.44928 | p-value=  0.449318 | p-value=  0.586712 | p-value=  0.524122 | p-value=  0.593633 |
| 50 | p-value=  0.466773 | p-value=  0.466783 | p-value=  0.282107 | p-value=  0.278325 | p-value=  0.270022 |
| 100 | p-value=  0.35684 | p-value=  0.356842 | p-value=  0.580793 | p-value=  0.513227 | p-value=  0.589205 |
| 200 | p-value=  0.00736531 | p-value=  0.00736518 | p-value=  0.0281456 | p-value=  0.0160219 | p-value=  0.0275328 |
| 500 | p-value=  0.000688338 | p-value=  0.000688332 | p-value=  0.0208024 | p-value=  0.0103168 | p-value=  0.0207784 |
| 1000 | p-value=  4.35605E-05 | p-value=  4.35602E-05 | p-value=  0.00130509 | p-value=  0.000352727 | p-value=  0.00130968 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется примерно *n* >160 (для критериев Бартлетта и Фишера) и *n* >170 (для остальных критериев).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется примерно *n* >170 (для критериев Бартлетта и Фишера) и *n* >190 (для остальных критериев).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется примерно *n* от 200 до 400 (для критериев Ансари-Бредли и Сижеля-Тьюки), *n* от 200 до 450 (для критерия Муда) и *n* > 200 (для критериев Бартлетта и Фишера).

3. *Смоделировать 2 выборки, одну − в соответствии со стандартным нормальным законом объемом =1000, другую − по логистическому закону с параметрами сдвига и масштаба (0, 0.551328895).*

*Последовательно, меняя объемы выборок   (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 .*

|  |  |  |
| --- | --- | --- |
| n | Критерий Смирнова | Критерий Лемана-Розенблатта |
| 50 | p-value=0.533 | p-value=0.45597 |
| 100 | p-value=0.97 | p-value=0.86219 |
| 500 | p-value=0.599 | p-value=0.702144 |
| 1000 | p-value=0.628 | p-value=0.454864 |
| 5000 | p-value=0.00431687 | p-value=0.00922302 |
| 10000 | p-value= 0.000373814 | p-value= 0.000244744 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется примерно *n* >3700 (для критерия Лемана-Розенблатта) и *n* >4300 (для критерия Смирнова).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется примерно *n* >4800.

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода потребуется примерно *n* >5000.