Министерство Образования и Науки Российской Федерации

Новосибирский Государственный Технический Университет

Кафедра теоретической и прикладной информатики

ЛАБОРАТОРНАЯ РАБОТА № 8

по предмету: История и методология прикладной математики и информатики

на тему «Применение критериев проверки гипотез о случайности и отсутствии тренда в данных»

#### 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, предназначенных для проверки статистических гипотез о случайности данных и отсутствии в анализируемых выборках признаков тренда. 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. *В предложенных 5 выборках объемом*  *первая половина моделировалась в соответствии с нормальным законом с параметрами сдвига и масштаба (0, 1), вторая половина* − *с параметрами (0.1, 1), то есть математическое ожидание для второй части (скачком) увеличилось на 10% от стандартного отклонения.*

*Последовательно, используя критерии автокорреляции, Кокса-Стьюарта и Бартелса, проверьте гипотезу о случайности и отсутствии тренда в 5 предложенных выборках. Проследите, как меняется достигаемый уровень значимости при проверке гипотезы.*

*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* | Кр. автокорреляции | Кр. Кокса-Стьюарта | Кр. Бартелса |
| 200 | p-value= 0.175076 | p-value= 0.932364 | p-value= 0.288329 |
| 400 | p-value= 0.146968 | p-value= 0.985343 | p-value= 0.696399 |
| 600 | p-value= 0.00842138 | p-value= 0.0336699 | p-value= 0.29522 |
| 800 | p-value= 0.0210284 | p-value= 0.151471 | p-value= 0.183246 |
| 1000 | p-value= 0.0198593 | p-value= 0.0230649 | p-value= 0.105145 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* >500 (для критерия автокорреляции), *n*  от 590 до 720 и n > 900 (для критерия Кокса-Стьюарта) и n > 1000 (для критерия Бартелса).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n >* 500(для критерия автокорреляции), *n*  от 600 до 680 и n > 980 (для критерия Кокса-Стьюарта).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* от 600 до 700 (для критерия автокорреляции), *n*  от 610 до 670 (для критерия Кокса-Стьюарта).

1. *Создайте (смоделируйте) аналогичные выборки, в которых в середине математическое ожидание изменяется на величину 30% от стандартного отклонения. Это можно сделать объединением 2-х выборок, смоделированных с различными параметрами сдвига.*

*Как и в п.1, используя критерии автокорреляции, Кокса-Стьюарта и Бартелса, проверьте гипотезу о случайности и отсутствии тренда в полученных выборках. Проследите, как меняется достигаемый уровень значимости при проверке гипотезы в данном случае. Насколько уменьшатся объемы выборок с обнаруженным трендом (с отклонением гипотезы о случайности).*

|  |  |  |  |
| --- | --- | --- | --- |
| *n* | Кр. автокорреляции | Кр. Кокса-Стьюарта | Кр. Бартелса |
| 200 | p-value= 0.157866 | p-value= 0.0523893 | p-value= 0.176906 |
| 400 | p-value= 0.00471185 | p-value= 0.142308 | p-value= 0.513611 |
| 600 | p-value= 0.0621666 | p-value= 0.085432 | p-value= 0.111735 |
| 800 | p-value= 0.496778 | p-value= 0.0169872 | p-value= 0.134162 |
| 1000 | p-value= 0.819299 | p-value= 0.00145303 | p-value= 0.126175 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* от 250 до 620 (для критерия автокорреляции), *n* < 300 и ** (для критерия Кокса-Стьюарта) и *n* от 610 до 700 (для критерия Бартелса).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* от 300 до 590 (для критерия автокорреляции), *n* > 700 (для критерия Кокса-Стьюарта).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* от 400 до 500 (для критерия автокорреляции), *n*  от 900 до 1000 (для критерия Кокса-Стьюарта).

1. *Смоделируйте (аналогично п. 2) 5 выборок*  *с наличием «сдвига в дисперсии», так чтобы в первой половине элементы моделировались в соответствии с нормальным законом с параметрами (0, 1), а во второй* − *с параметрами (0, 1.1), то есть стандартное отклонение второй части отличается на 10% от первой в большую сторону.*

*Последовательно, используя критерий Хсу с h-статистикой, критерий Кокса-Стьюарта для дисперсий, ранговые критерии с метками Клотца и Сэвиджа, проверьте гипотезу о случайности и отсутствии тренда в 5 построенных выборках. Проследите, как меняется достигаемый уровень значимости при проверке гипотезы.*

*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* | Кр. Xcy с h-статистикой | Кр. Кокса-Стьюарта для дисперсий | Ранговый критерий с метками Клотца | Ранговый критерий с метками Сэвиджа |
| 200 | p-value= 0.989239 | p-value= 0.181354 | p-value= 0.858334 | p-value= 0.664933 |
| 400 | p-value= 0.675361 | p-value= 0.918196 | p-value= 0.886981 | p-value= 0.291064 |
| 600 | p-value= 0.130354 | p-value= 0.709378 | p-value= 0.242608 | p-value= 0.652378 |
| 800 | p-value= 0.35638 | p-value= 0.656029 | p-value= 0.586045 | p-value= 0.268465 |
| 1000 | p-value= 0.0688911 | p-value= 0.0896152 | p-value= 0.542844 | p-value= 0.204673 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* >1000 (для критериев Xcy с h-статистикой и Кокса-Стьюарта).

4. *Аналогично п. 2 создайте (смоделируйте) аналогичные выборки, в которых в середине стандартное отклонение (параметр масштаба) изменяется в большую сторону на величину 30%. Как и ранее это можно сделать объединением 2-х выборок, смоделированных с различными параметрами масштаба.*

*Используя те же критерии, что и в п. 3 (критерий Хсу с h-статистикой, критерий Кокса-Стьюарта для дисперсий, ранговые критерии с метками Клотца и Сэвиджа), проверьте гипотезу о случайности и отсутствии тренда в 5 построенных выборках. Проследите, как меняется достигаемый уровень значимости при проверке гипотезы.*

*Зафиксируйте результаты проверок в таблице.*

*Оцените, насколько раньше (при каком объеме выборок) будет принято верное решение об отклонении «несправедливой» проверяемой гипотезы при задании вероятности ошибки 1-го рода .*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *n* | Кр. Xcy с h-статистикой | Кр. Кокса-Стьюарта для дисперсий | Ранговый критерий с метками Клотца | Ранговый критерий с метками Сэвиджа |
| 200 | p-value= 0.000425813 | p-value= 0.0419564 | p-value= 0.864239 | p-value= 0.156364 |
| 400 | p-value= 0.0632871 | p-value= 0.0433953 | p-value= 0.26814 | p-value= 0.00645205 |
| 600 | p-value= 0.000197095 | p-value= 0.112362 | p-value= 0.918227 | p-value= 0.00557783 |
| 800 | p-value= 0.000767828 | p-value= 0.130918 | p-value= 0.467247 | p-value= 0.0684655 |
| 1000 | p-value= 1.97318E-6 | p-value= 0.000138653 | p-value= 0.886784 | p-value= 0.0137238 |

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* от 200 до 550 (для критерия Кокса-Стьюарта), *n >* 350 (для рангового критерия с метками Сэвиджа).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n* от 200 до 400 (для критерия Кокса-Стьюарта), *n* от 200 до 300 и *n* >500 (для критерия Xcy с h-сатистикой) и *n* от 400 до 600 (для рангового критерия с метками Сэвиджа).

Чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу при задании вероятности ошибки 1-го рода , потребуется примерно *n=*1000 (для критерия Кокса-Стьюарта), *n* от 200 до 220 и *n* >600 (для критерия Xcy с h-сатистикой) и *n* от 400 до 600 (для рангового критерия с метками Сэвиджа).