Исследование мощности перестановочных тестов на основе статистического моделирования

Гудулина Анастасия Олеговна, гр. 522

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: д.ф.-м.н. проф. В.Б. Мелас
Рецензент: к.ф.-м.н. доц. П.В. Шпилев

СтатМод

Санкт-Петербург
2014г.

Введение

Задача сравнения двух распределений:

- классический t-тест обладает рядом оптимальных свойств;
- непараметрические критерии, важным классом которых являются критерии, основанные на перестановках.

Недавно такие критерии были исследованы в работе [Sirsky M., On the Statistical Analysis of Functional Data Arising from Designed Experiments, 2012] с помощью статистического моделирования для задачи дискриминации двух регрессионных моделей.

Исследование проводилось для ряда конкретных регрессионных моделей. В качестве модели ошибок рассматривался процесс авторегрессии первого порядка.

Введение

Основные выводы работы [Sirsky M., On the Statistical Analysis of Functional Data Arising from Designed Experiments, 2012]:

- перестановочные методы являются более мощными, чем параметрические методы;
- метод, основанный на попарном сравнении абсолютных разностей, является наилучшим из перестановочных методов.

Кроме того, в работе научного руководителя с итальянскими коллегами [Corain L., Melas V., Salmaso L. et al., 2013] была доказана эквивалентность некоторых перестановочных критериев, рассмотренных в работе [Sirsky M., 2012].

Постановка задачи

Проверка гипотезы

$$H_0 : F_1 = F_2$$

против альтернативы

$$H_1 : F_1 \neq F_2,$$

где F_1 и F_2 – два неизвестных распределения, на основе выборок Y_{ij}, $i = 1, 2; j = 1, \ldots, n_i$, соответственно из первого и второго распределений.

Цель дипломной работы: исследовать мощность четырех перестановочных критериев, выбранных на основе анализа результатов работы Sirsky, и классического t-критерия в зависимости от вида распределения и размера выборки.

Перестановочные критерии

$$Z(\pi_0) = \{Y_{11}, \ldots, Y_{1n}, Y_{21}, \ldots, Y_{2n}\},$$
$$Z(\pi_k) = \{\tilde{Y}_{11}, \ldots, \tilde{Y}_{1n}, \tilde{Y}_{21}, \ldots, \tilde{Y}_{2n}\},$$

$$\tilde{Y}_{1i} = Y_{2j},$$
$$\tilde{Y}_{2i} = Y_{1j},$$
$$\tilde{Y}_{1j} = Y_{1j}, j \neq i_1, ..., i_k,$$
$$\tilde{Y}_{2j} = Y_{2j}, j \neq j_1, ..., j_k,$$

где $\pi_k = \pi_k(s), s = 1, 2, \ldots, (C_n^k)^2$ - различные способы замены k элементов из первой половины на k элементов из второй половины.

$$K_1(Z) = (\overline{Y_1} - \overline{Y_2})^2,$$
$$K_4(Z) = (Y_{1med} - Y_{2med})^2,$$
$$K_5(Z) = (\sum_{i=1}^{n} |Y_{1i} - Y_{1med}| + \sum_{i=1}^{n} |Y_{2i} - Y_{2med}|)^2,$$
$$K_6(Z) = \sum_{i,j=1}^{n} |Y_{1i} - Y_{2j}|.$$

Перестановочные критерии

Перестановочный K_i-критерий проверки гипотезы H_0:

- пусть $r_2 = (C_n^{\frac{n}{2}})^2$, n – четное и пусть r_1 – число перестановок π, для которых $K_i\left(Z(\pi)\right) > K_i\left(Z(\pi_0)\right)$;
- если $\frac{r_1}{r_2} \geq \alpha$, где α – заданный уровень значимости, то нулевая гипотеза не отвергается при заданном α-уровне для критериев K_1, K_4 и K_6, для K_5 при $\frac{r_1}{r_2} \leq 1 - \alpha$;
- если $\frac{r_1}{r_2} < \alpha$, то нулевая гипотеза отвергается для критериев K_1, K_4 и K_6, для K_5 при $\frac{r_1}{r_2} > 1 - \alpha$.

Распределения

F_1:

- нормальное распределение $N(0, 1)$
- распределение Коши $C(0, 1)$
- смесь 95% нормального распределения $N(0, 1)$ и 5% распределения Коши $C(0, 1)$
- равномерное распределение $U(0, 1)$
- распределение Вейбулла $W(1, 3)$

F_2:

- нормальное распределение $N(\mu, 1)$, $\mu = 0, 1, 1.5, 2$
- распределение Коши $C(x_0, 1)$, $x_0 = 0, 5, 10, 15, 20$
- смесь 95% нормального распределения $N(\mu, 1)$, $\mu = 0, 1, 1.5, 2$ и 5% распределения Коши $C(0, 1)$
- равномерное распределение $U(0 + shift, 1 + shift)$, где $shift = 0, 0.2, 0.4, 0.6, 0.8$
- распределение Вейбулла $W(1, b)$, $b = 3, 2.5, 2, 1.5, 1, 0.5$

Число перестановок

- Всего возможно $(C_n^{\frac{n}{2}})^2$ вариантов замены одной половины выборки на другую.
- При большом n число перестановок становится неимоверно большим. В таких случаях мы можем брать случайную выборку перестановок.
- Критерий перестановок, основанный на подмножестве из 1600 перестановок, имеет мощность, близкую к мощности теста с использованием всех перестановок.

Мощность тестов

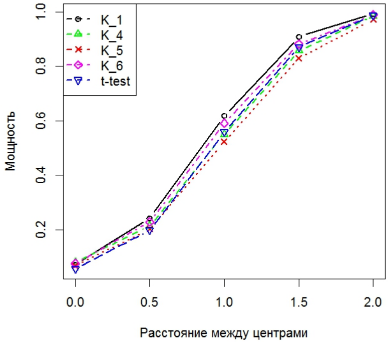

Рис. 1: Нормальное распределение, $n = 10$

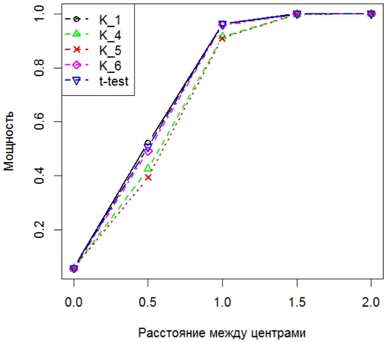

Рис. 2: Нормальное распределение, $n = 30$

Мощность тестов

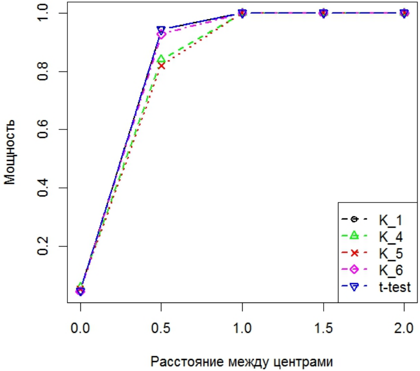

Рис. 3: Нормальное распределение, $n = 100$

Мощность тестов

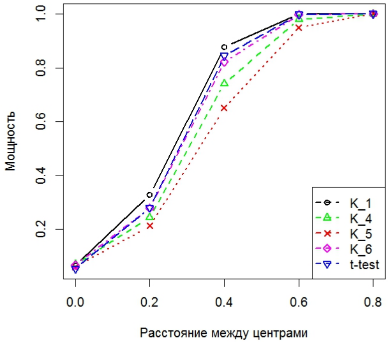

Рис. 4: Равномерное распределение, $n = 10$

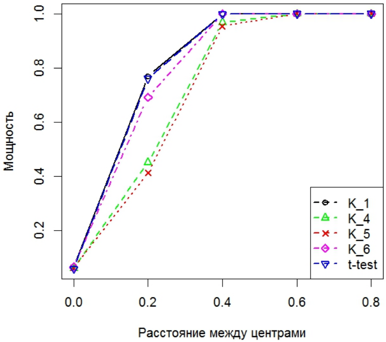

Рис. 5: Равномерное распределение, $n = 30$

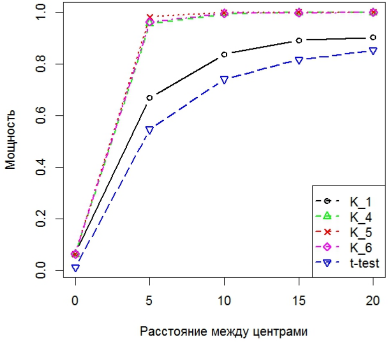

Рис. 6: Распределение Коши, $n = 10$

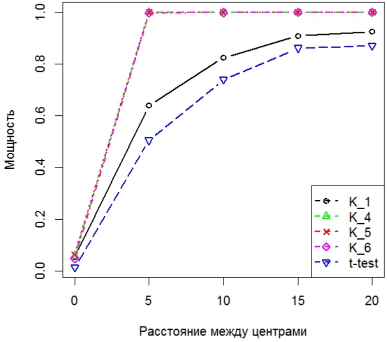

Рис. 7: Распределение Коши, $n = 30$

Мощность тестов

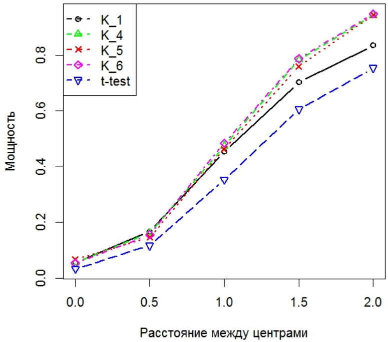

Рис. 8: Смесь нормального распределения и Коши, $n = 10$

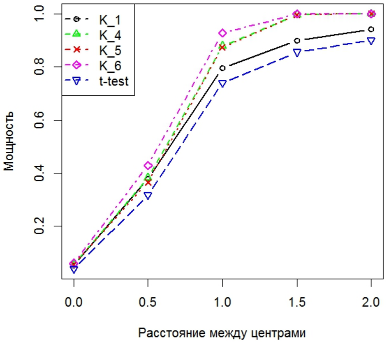

Рис. 9: Смесь нормального распределения и Коши, $n = 30$

Мощность тестов

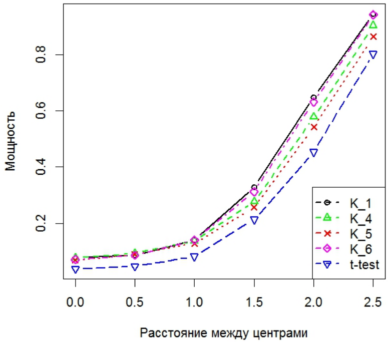

Рис. 10: Распределение Вейбулла, $n = 10$

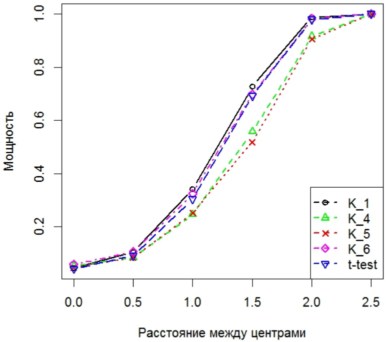

Рис. 11: Распределение Вейбулла, $n = 30$

Мощность тестов

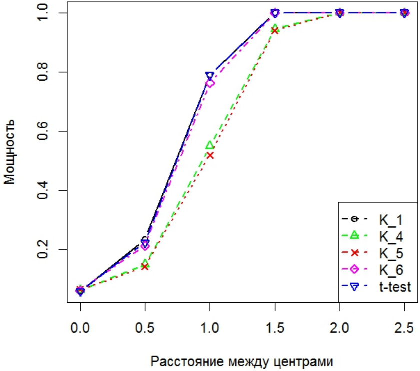

Рис. 12: Распределение Вейбулла, $n = 100$

Мощность тестов

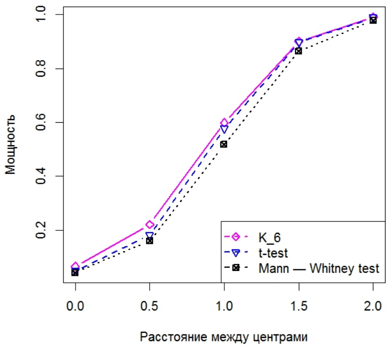

Рис. 13: Нормальное распределение, $n = 10$

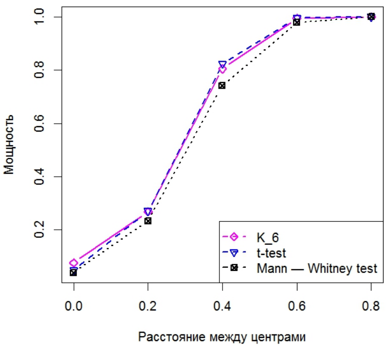

Рис. 14: Равномерное распределение, $n = 10$

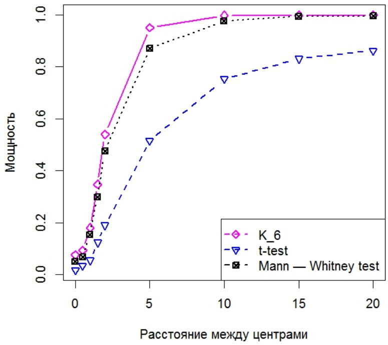

Рис. 15: Распределение Коши, $n = 10$

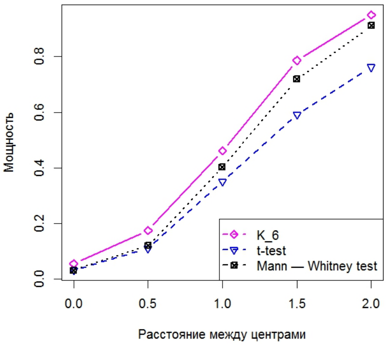

Рис. 16: Смесь нормального распределения и Коши, $n = 10$

Мощность тестов

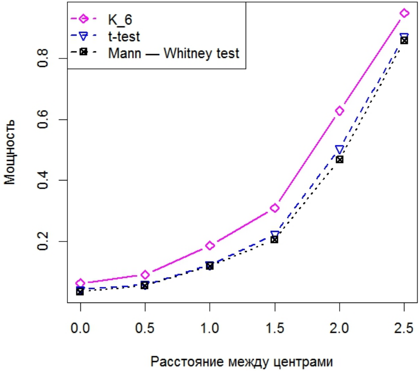

Рис. 17: Распределение Вейбулла, $n = 10$

Итоги

- Перестановочный тест K_6, основанный на попарном сравнении абсолютных величин разностей, оказывается либо наиболее мощным, либо очень близок к наиболее мощному из рассмотренных тестов.
- Для малых выборок t-тест уступает одному или даже всем перестановочным тестам при любом из рассмотренных распределений.