Сравнение параметрических и непараметрических тестов с помощью статистического моделирования

Сальников Дмитрий Игоревич, гр. 422

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

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

Санкт-Петербург
2016г.

Преследуемые цели:

- Выяснить, эффективно ли использовать перестановочные тесты для проверки статистических гипотез.
- Сравнить их мощности с мощностями наиболее популярных неперестановочных тестов в зависимости от вида распределения и размера выборки.
- Дать общие рекомендации для тех случаев, когда эффективно применять тот или иной тест.

Две выборки

$$X_{ij} \sim F_i,\ i = 1, 2,\ j = 1 \dots n_i,$$

нулевая гипотеза

$$H_0 : F_1 = F_2,$$

альтернативная гипотеза

$$H_1 : F_1 \neq F_2.$$

Будем считать $n_1 = n_2 = n; \quad n = \{10, 30, 100\}.$

Постановка задачи

Исследуемые тесты:

- Перестановочные тесты K_1, K_2, K_3, K_4, K_5, K_6,
- тест Стьюдента,
- тест Колмогорова-Смирнова,
- тест Манна-Уитни.

Исходная объединенная выборка

$$Z(\pi_0) = X_{11}, \ldots, X_{1n}, X_{21}, \ldots, X_{2n},$$

перестановки

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

где $\pi_k, \ k = 0 \ldots n$ — различные способы замены k элементов первой выборки на k элементов второй.

$K_i(Z)$ — статистика перестановочного теста K_i.

Перестановочные тесты

- $K_1(Z) = (\overline{X}_1 - \overline{X}_2)^2$
  (Sturino J., et al., 2010),
- $K_2(Z) = \sum_{i,j=1}^{n} (X_{1i} - X_{2j})^2 / n^2$
  (Sirsky M., 2012),
- $K_3(Z) = nK_1(Z)/(S_1^2(Z) + S_2^2(Z))$,
  $S_i^2(Z) = \frac{1}{n} \sum_{j=1}^{n} (X_{ij} - \overline{X}_i)^2$
  (Cox D., Lee J., 2008 and Ramsay J., et al., 2009).

Теорема (Melas V. et al., 2013)

Перестановочные тесты K_1, K_2 и K_3 для проверки гипотезы однородности эквивалентны для любой перестановки и для любого произвольно заданного уровня значимости α.

Рассмотренные перестановочные тесты:

- $K_1(Z) = (\overline{X}_1 - \overline{X}_2)^2$
  (Sturino J., et al., 2010),
- $K_4(Z) = (X_{1med} - X_{2med})^2$
  (Sirsky M., 2012),
- $K_5(Z) = (\sum_{i=1}^{n} |X_{1i} - X_{1med}| + \sum_{i=1}^{n} |X_{2i} - X_{2med}|)^2$,
- $K_6(Z) = \sum_{i,j=1}^{n} |X_{1i} - X_{2j}|$
  (Sirsky M., 2012).

Перестановочные тесты

Перестановочный K_i-тест проверки гипотезы H_0:

- пусть r_2 — общее число перестановок, r_1 — число перестановок π_k, для которых $K_i(Z(\pi_k)) > K_i(Z(\pi_0))$, $\alpha = 0.05$ — заданный уровень значимости;
- H_0 не отвергается при α-уровне значимости для тестов K_1, K_4, K_6, если $\frac{r_1}{r_2} \geq \alpha$, для K_5 — если $\frac{r_1}{r_2} \leq 1 - \alpha$.

При моделировании $r_2 = 1600$, перестановки случайны (Keller-McNulty S., Higgins J., 1987).

Распределения

- Нормальное распределение N(μ, σ)
- Распределение Коши C(x_0, γ)
- Смесь 95% N(μ, σ) и 5% C(0, 1)
- Распределение Стьюдента t(n, x_0)
- Распределение Фишера F(d_1, d_2)
- Бета-распределение B(α, β)
- Гамма-распределение G(k, θ)
- Равномерное распределение U(a, b)
- Распределение Вейбулла W(k, λ)

Вспомогательное предложение

Предложение

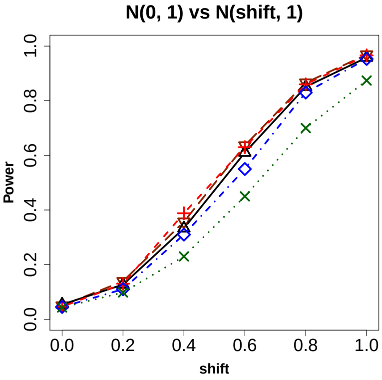
Пусть в N экспериментах получена оценка мощности p_1 по первому тесту и p_2 по второму. Тогда, если N достаточно велико и

$$p_2 > p_1 + 3\frac{p_1(1-p_1) + p_2(1-p_2)}{\sqrt{N}},$$

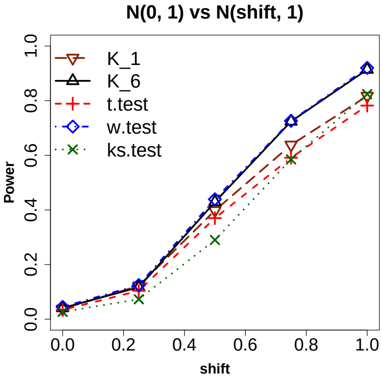
то с вероятностью более чем 99% второй тест является более мощным, чем первый.

Второй тест мощнее первого с вероятностью более 99%

Численные результаты в графической форме

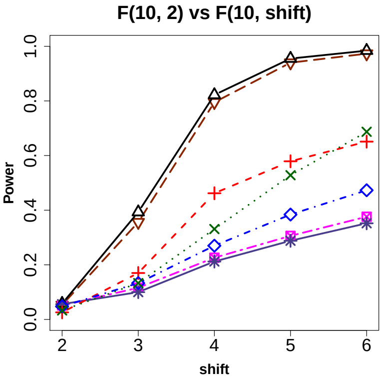


Нормальное распределение

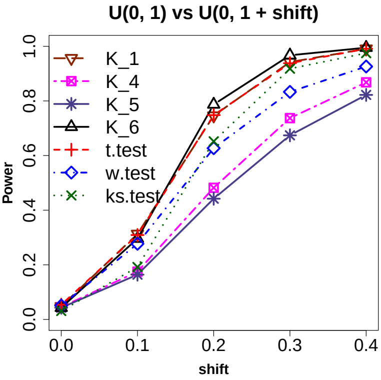


Загрязненное нормальное распределение

Численные результаты в графической форме

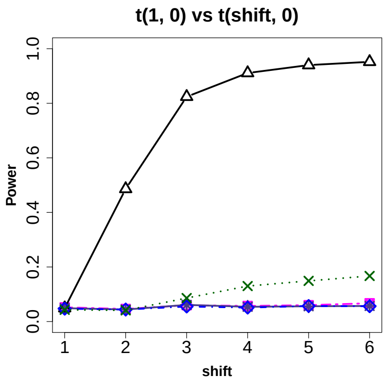


Распределение Фишера

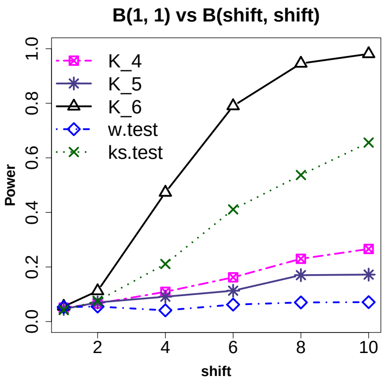


Равномерное распределение

Численные результаты в графической форме

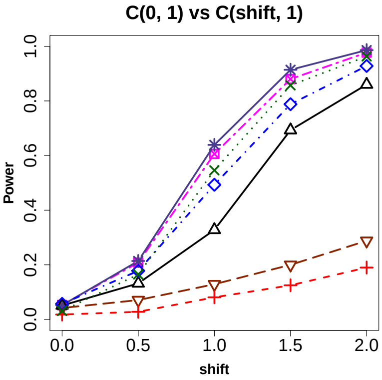


Распределение Стьюдента

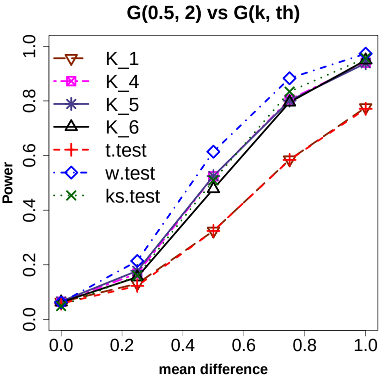


Бета-распределение

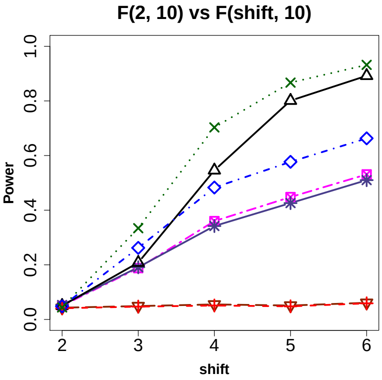
Численные результаты в графической форме



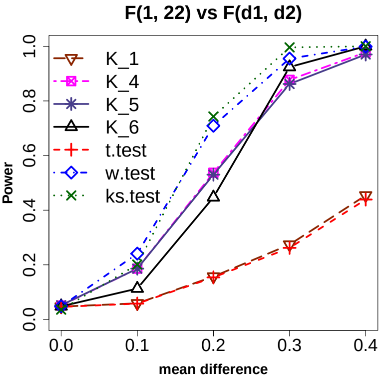
Распределение Коши



Гамма-распределение



Распределение Фишера



Распределение Фишера

Итоги

- Тест K_6 является наиболее мощным среди всех рассмотренных тестов, за исключением ряда случаев. Особенно велико преимущество этого теста, если распределения симметричны относительно общего центра.
- В случаях сдвига распределений Коши, Фишера и Гамма и при изменении первого параметра распределения Фишера K_6 значительно уступает лидирующим в мощности тестам.
- Тесты K_4 и K_5 эффективны в случае сдвига распределения Коши, в других случаях уступают в мощности K_6.
- По результатам работы подготовлена статья, принятая к печати в журнале «Вестник СПбГУ, сер. 1» вып. 3 (2016).