

Перестановочные тесты и бутстреп анализ

Подход с перестановками

- Вычисляют t-статистику как при параметрическом подходе и получают статистику t_0
- Объединяют все объекты в одну группу
- Делят на группы каждый раз помещая в них новые сочетания объектов, каждый раз вычисляя новую t-статистику (для 10 объектов – 252 способа перестановки)
- Располагают все значения t-статистики по возрастанию, получая эмпирическое распределение, основанное на выборках
- Если t_0 не входит в центральные 95% значений эмпирического распределения, отвергают H_0 о равенстве средних в двух группах

Перестановочные тесты из пакета coin

Перестановочный тест

Функция

Тест для двух и k выборок (аналог теста Стьюдента и однофакторного дисперсионного анализа)

`oneway_test(y~A)`

Тест Вилкоксона для независимых выборок (тест Манна-Уитни)

`wilcox_test(y~A)`

Тест Вилкоксона для зависимых выборок

`wilcoxsign_test(y1~y2)`

Chi-квадрат тест Пирсона

`chisq_test(A~B)`

Тест Кохрана-Мантеля-Хензеля

`cmh_test(A~B | C)`

Тест Спирмена

`spearman_test(y~x)`

Линейные модели: дисперсионный и регрессионный анализ

Пакет ImPerm

Регрессионный анализ: lmp()

Дисперсионный анализ: aovp()

Бутстреп-анализ

- Создает эмпирическое распределение тестовой статистики путем создания многих случайных выборок, основанных на исходной выборке
- Позволяет вычислять доверительные интервалы и проверять статистические гипотезы без опоры на определенное теоретическое распределение

Бутстреп-анализ

Имеем 10 наблюдений:

Нужно определить *доверительный интервал*

1. Случайно выбрать 10 наблюдений из выборки с возвратом значений после каждого выбора
2. Вычислить среднее для полученной выборки
3. Повторить шаги 1-2 (тысячу) раз
4. Отсортировать тысячу выборочных средних по возрастанию
5. Найти выборочные средние, которые представляют собой 2,5 и 97,5 процентиля. Это и будут границы 95%-го доверительного интервала

Пакет boot

Установить пакет “boot”

- 3 этапа бутстреп-анализа
1. Написать функцию, которая вычисляет нужную статистику.
 2. Применить функцию `boot()` к этой функции, чтобы создать бутстреп-повторности данной статистики
 3. Использовать функцию `boot.ci()`, чтобы вычислить доверительные интервалы для искомой статистики